Ontwerp van geavanceerde benchmarks en analytische methodes voor RF-gebaseerde indoorlokalisatie-oplossingen

Design of Advanced Benchmarks and Analytical Methods for RF-Based Indoor Localization Solutions

Tom Van Haute

Promotoren: prof. dr. ir. E. De Poorter, prof. dr. ir. I. Moerman
Proefschrift ingediend tot het behalen van de graad van
Doctor in de ingenieurswetenschappen: computerwetenschappen

UNIVERSITEIT GENT

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. D. De Zutter
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2016 - 2017

ISBN 978-90-8578-936-9
NUR 986, 975
Wettelijk depot: D/2016/10.500/68

UNIVERSITEIT GENT

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur
Vakgroep Informatietechnologie

Promotoren: prof. dr. ir. Eli De Poorter
prof. dr. ir. Ingrid Moerman

Leden van de Examencommissie:

prof. Luc Taerwe (Voorzitter)
prof. Wout Joseph (Secretaris)
prof. Eli De Poorter (Promotor)
prof. Ingrid Moerman (Promotor)
prof. Guido Dolmans (TU Eindhoven, Nederland)
prof. Maarten Weyn (Universiteit Antwerpen)
prof. Eric Laermans (UGent, INTEC)
prof. Heidi Steendam (UGent, TELIN)

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur

Vakgroep Informatietechnologie
iGent Tower, Technologiepark 15, 9052 Gent (Zwijnaarde), België

Tel.: +32-9-331.49.00
Fax.: +32-9-331.48.99

INTEC

Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen:
Computerwetenschappen
Academiejaar 2016-2017

Dankwoord

Geloof het of niet, maar dit deel was moeilijk voor me om te schrijven, om niet te zeggen, hét moeilijkste. Dit boek mag dan wel mijn naam dragen, er zijn veel mensen aan te pas gekomen die mee verantwoordelijk zijn voor de realisatie van dit werk. Het is dan ook een onmogelijke opdracht iedereen te bedanken die heeft bijgedragen aan dit onderzoek. Toch wil ik een dappere poging ondernemen in dit hoofdstuk.

In de eerste plaats wil ik graag mijn promotoren, prof. Eli De Poorter en prof. Ingrid Moerman, bedanken voor de gekregen kansen, de continue ondersteuning, raad en bijstand in het volledige traject. Graag wil ik Eli in het bijzonder bedanken voor zijn grondige paper reviews en kritische feedback op al mijn werk. Verder wil ik ook prof. Piet Demeester bedanken dat ik deel mocht uitmaken van de onderzoeksgroep IBCN.

Ook gaat mijn welgemeende dank uit naar alle andere leden van de examencommissie, naast mijn promotoren, voor het grondig evalueren van mijn werk, de interessante vragen en discussies. Hartelijk bedankt prof. Luc Taerwe, prof. Guido Dolmans, prof. Maarten Weyn, prof. Wout Joseph, prof. Eric Laermans en prof. Heidi Steendam.

Bij een doctoraat komen er heel wat administratieve taken bij kijken. Hartelijk dank Martine en Davinia voor al mijn paperassen, Joke en Bernadette voor mijn onkostennota's, de admins voor de server en alle kabeltjes die ik iedere keer weer kwam vragen, en Sabrina voor de aangename babbels en de propere werkomgeving.

Heel veel dank voor Bart, Pieter en Vincent, onze testbed admins. Ik denk dat ik één van de meest veeleisende onderzoekers was voor w-iLab.t II, vooral wat de robots betreft. Bedankt voor de snelle support, elke keer opnieuw.

Verder had ik de eer en het genoegen om samen met Daan, David, Dries, Elnaz, Floris, Girum, Irina, Isam, Jen, Michael, Pieter, Tarik, Wei en Wim in bureau 3.17 te vertoeven. Bedankt voor de leuke sfeer en interessante babbels! Jullie zorgde ervoor dat ik elke dag met veel plezier naar de Zuiderpoort afzakte. Na onze verhuis naar iGent mag het lijstje worden aangevuld met Jetmir en Vasileios. Spijtig genoeg was ik daar niet meer zoveel aanwezig, maar toch ook een welgemeende dankjewel.

En dan zijn er uiteraard nog veel mensen die buiten UGent ook een belangrijke rol hebben gespeeld. Bedankt Frederik, Jochen, Tom en Willem, collega's van Fleetmaster, waar ik toch ook een periode zeer intensief mee heb samen gewerkt. Veel dank ook voor de voltallige Hela-crew, de batterijen werden telkens weer goed opgeladen iedere keer we samen kwamen. De zeldzame tijd die me dan nog restte, kon ik altijd opvullen met 'de boys' (vergezeld met al enkele 'girls' ondertussen). Het waren fijne tijden: de avonden, de feestjes, de roadtrips en alle andere momenten. Dat er zo nog veel mogen volgen! Dikke merci Axel, Celine, Frederik, Jan, Sofie, Steven, Tina, Verm en Willem!

Dan rest er mij nog enkel de belangrijkste personen te bedanken. Bedankt lieve mama en papa voor alle steun en kansen die ik gekregen heb. Jullie positieve stimulatie en interesse hebben ervoor gezorgd dat ik de eindmeet heb gehaald. Merci ook superzus Tineke dat je *apetrots* bent op je broertje! Tot slot, nog een ongelooflijk woord van dank voor mijn steun en toeverlaat, mijn vriendin Evi. Bedankt voor de motivatie, bedankt dat ik zowel mijn euforie als mijn frustraties met jou kon delen, bedankt om er altijd te zijn, bedankt voor alles!

Sint-Niklaas, september 2016
Tom Van Haute

Table of Contents

List of Figures

List of Tables

List of Acronyms

0-9

5G 5th Generation

A

AAL Ambient Assisted Living
ACK Acknowledgement
AoA Angle of Arrival
AP Access Point
API Application Programming Interface

B

BLE Bluetooth Low Energy

C

CDF Cummulative Distribution Function
COTS Commercial Off-The-Shelf
CREW Cognitive Radio Experimentation World

D

DECT Digital Enhanced Cordless Telecommunications

DToA Differential Time of Arrival

E

EBH EVARILOS Benchmarking Handbook
EBP EVARILOS Benchmarking Platform
EBS EVARILOS Benchmarking Suite
EC European Commission
ED Euclidean Distance
EE Environment Emulator
EvAAL Evaluating AAL Systems through Competitive Benchmarking
EVARILOS Evaluation of RF-based Indoor localization Solutions

F

FIRE Future Internet Research & Experimentation
FP7 7^{th} Framework Programme for Research and Technological Development

G

GNSS Global Navigation Satellite System
GPS Global Positioning System
GSM Global System for Mobile communications

H

HTTP Hypertext Transfer Protocol

I

IEC International Electrotechnical Commission

IEEE Institute of Electrical and Electronics Engineers
ILS Indoor Localization Sensing
IMU Inertial Measurement Unit
IPIN Indoor Positioning & Indoor Navigation
IPS Indoor Positioning System
IPSN Information Processing in Sensor Networks
IR Infra-Red
ISM Industrial, Scientific and Medical
ISO International Organisation for Standardization

J

JTC Joint Technical Commitée

K

KL Kullback-Leibler
kNN k-Nearest Neighbours

L

LBS Location-Based Service
LoS Line-of-Sight
LQI Link Quality Indicator

M

MLAT Multilateration
MSE Mean-Square Error
MvG Multivariate Gaussian

N

NFC Near Field Communication
NIC Network Interface Card
NLoS Non Line of Sight

O

OMF cOntrol and Management Framework

P

PDF Probability Density Function
PH Pompeiu-Hausdorff

R

REST REpresentational State Transfer
RF Radio-Frequency
RFID Radio-Frequency Identification
RMS Root-Mean Square
ROI Return On Investment
ROS Robot Operating System
RSS Received Signal Strength
RSSI Received Signal Strength Indicator
RTT Round Trip Time

S

SDK Software Development Kit
SDS-TWR Symmetrical Double-Sided Two-Way Ranging
SSID Service Set Identifier

SUT	System Under Test

T

TDoA	Time-Difference of Arrival
TKN	Telekommunikationsnetze (Telecommunication Networks Group)
ToA	Time of Arrival
ToF	Time of Flight
TUB	Technische Universität Berlin
TWIST	TKN Wireless Indoor Sensor network Testbed
TX	Transmission power

U

US	Ultrasound
UWB	Ultra-Wideband

W

WS	Weighted Sum
WSN	Wireless Sensor Networks

Samenvatting
– Summary in Dutch –

In de laatste jaren is locatie-gebaseerde informatie onontbeerlijk geworden voor meerdere applicatiedomeinen. Gezien Line-of-Sight (LoS) een vereiste is voor Global Positioning System (GPS)-technologieën, is deze beperkt tot omgevingen buitenshuis. Het feit dat draadloze technologieën (zoals Wi-Fi, Bluetooth Low Energy (BLE), etc.) populairder worden in onze omgeving, schept mogelijkheden: Indoor Positioning Systems (IPSs). In de meeste systemen kunnen twee soorten nodes worden onderscheiden. Ten eerste, anker nodes worden op tactische locaties in het gebouw geïnstalleerd. Deze zullen zich gedragen als de referentiepunten van het systeem. Vervolgens, de persoon of het object dat gelokaliseerd moet worden, draagt of is voorzien van een mobiele node. Het IPS zal trachten de coördinaten van deze mobiele node te bepalen. Tegenwoordig kunnen verschillende technologieën en technieken worden onderscheiden. Om die reden stijgt het aantal indoor lokalisatie toepassingen aanzienlijk en moeten we bewust zijn van het feit dat een meer geavanceerde manier van evaluatie en vergelijken noodzakelijk is.

Jammer genoeg is er momenteel nog geen gestandaardiseerde evaluatiemethode om deze grote variatie aan oplossingen te vergelijken. Dit heeft tot gevolg dat elke oplossing is geëvalueerd onder individuele, niet vergelijkbare en niet herhaalbare condities. Bovendien is de selectie van evaluatiepunten vaak bevooroordeeld: de best bereikbare evaluatiepunten worden geselecteerd. Een ander probleem situeert zich in het rapporteren van de bekomen resultaten: de meerderheid van de evaluaties focussen zich hoofdzakelijk op punt accuraatheid terwijl de andere cruciale metrieken worden gegenereerd zoals reactietijd, energieverbruik, schaalbaarheid, kost, eenvoud, etc. Bovendien worden deze resultaten vaak gepresenteerd met behulp van statistische waarden zoals gemiddelde of mediaan. Deze manier van evalueren en rapporteren maakt het onmogelijk om objectief verschillende oplossingen met elkaar te vergelijken.

Deze dissertatie pakt de bovenvermelde tekortkomingen aan door een vernieuwde benchmarking methodologie voor te stellen. Een handboek definieert verschillende metrieken en benchmarking scenario's om een IPS op een objectieve manier te evalueren. Vervolgens is er een web-gebaseerd platform ter beschikking gesteld dat is afgestemd met het handboek. Het beschikt over verschillende tools om de onderzoeker te ondersteunen bij de evaluatie van zijn oplossing. Dit platform geeft de mogelijkheid om (i) datasets te maken en te delen, (ii) geschatte coördinaten te uploaden, (iii) prestatie resultaten te visualiseren en te vergelijken om uiteindelijk (iv) scores te berekenen en rangschikken gebaseerd op een zelf gedefinieerd applicatie profiel.

Verder is het belangrijk om een onderscheid te maken tussen een "lokalisatie" en een "tracking" systeem. Beide systemen proberen om de locatie van de mobiele node in te schatten. Maar de laatstgenoemde maakt gebruik van informatie gebaseerd op de voorgaande (geschatte) posities van de mobile node. Hierdoor kan een volledig pad worden gereconstrueerd. Deze informatie is cruciaal voor dit type systemen om de locatie van de mobile node te kunnen bepalen. Gezien hierdoor de mobile node continu moet communiceren om zijn locatie up-to-date te houden, beperkt deze dissertatie zich tot indoor *lokalisatie* toepassingen.

Initieel start dit werk met het opsporen van problemen wanneer een Time of Arrival (ToA)-gebaseerde oplossing wordt toegepast in een complexe industriële omgeving. Gezien deze omgeving een open ruimte is met veel metalen objecten, veroorzaken draadloze signalen "multipath fading" door de reflecties. Multipath fading is een van de meest voorkomende redenen waarom incorrecte metingen (van Received Signal Strength Indicator (RSSI), ToA, etc.) worden verzameld. Geavanceerde kalibratie en filtertechnieken kunnen deze incorrecte metingen detecteren en elimineren. Het filteren van deze incorrecte metingen verbetert de oplossing zijn prestatie en versterkt de robuustheid aanzienlijk.

Het tweede gedeelte van deze dissertatie stelt een grondige vergelijking en analyse voor tussen verschillende IPSs in gevarieerde omgevingen binnenshuis. Drie verschillende oplossingen zijn geanalyseerd: een IEEE 802.15.4 gebaseerd op ToA, een IEEE 802.15.4 gebaseerd op RSSI multilateratie en ten slotte een IEEE 802.11 oplossing gebaseerd op fingerprinting. Deze zijn geëvalueerd in drie verschillende industriële omgevingen (elk met een ander type muur). Alle evaluaties zijn uitgevoerd met dezelfde methode. Het resultaat van dit werk legde enkele problemen bloot: (i) er bestaan meerdere inherente compromissen tussen verschillende metrieken, die meestal genegeerd worden door enkel de punt accuraatheid van de oplossing te publiceren. (ii) De resultaten betrekkende tot de accuraatheid zijn sterk afhankelijk van de karakteristieken van de omgeving. (iii) Ten slotte, de keuze van de evaluatiepunten hebben een sterke invloed op de gerapporteerde accuraatheid.

IPSs zijn vaak geëvalueerd in "geoptimaliseerde" omgevingen zonder bijkomstige (gegenereerde) interferentie. Dit is een andere manier om de resultaten te verbloemen. Om dit na te gaan, is in deze dissertatie een analyse voorgesteld van de bovenvermelde oplossingen in een ziekenhuisomgeving waarbij medische apparatuur realistische interferentie genereerde gezien het ziekenhuis actief in gebruik was. Dit werk bevestigt de reeds vermelde bevindingen. Dit werk gaat nog een stap verder door de invloed van de anker nodes na te gaan. De robuustheid van de oplossingen is onderworpen aan een stress-test door een of meerdere anker nodes uit te schakelen.

Het belangrijkste gedeelte van dit onderzoek was het definiëren van een benchmarking methodologie om ervoor te zorgen dat verschillende oplossingen met elkaar vergeleken konden worden. Het startpunt was het ontwerp van een benchmarking handboek. Een handboek waarin procedures zijn beschreven die het mogelijk maken om oplossingen objectief te vergelijken. Deze methodologie is afgestemd op de opkomende ISO/IEC 18305 standaard "Test en evaluatie van lokalisatie en tracking systemen". Verder beschrijft het handboek een set gevalideerde, gestandaardiseerde en experiment gebaseerde scenario's die gefocust zijn op de evaluatie van Radio-Frequency (RF) gebaseerde lokalisatie oplossingen. Deze worden aangevuld met duidelijke definities en voorbeelden van de prestatie metrieken zoals punt & kamer accuraatheid, reactietijd, energieverbruik, installatietijd, herhaalbaarheid, gevoeligheid aan interferentie en omgeving, etc.

Vervolgens wordt een web-gebaseerd platform voorgesteld. Dit platform bestaat uit verschillende tools die de onderzoeker ondersteunen om een objectief resultaat van zijn oplossing te bekomen. Elke gebruiker kan nieuwe datasets aanmaken gebaseerd op een persoonlijke configuratie. Accurate Roomba gebaseerde robots worden ingeschakeld om de datasets te verkrijgen in de w-iLab.t II testomgeving. Op deze manier wordt herhaalbaarheid mogelijk waardoor de robuustheid van de oplossing onder verschillende condities kan worden geëvalueerd. Deze datasets (die RSSI of ToA metingen bevatten) zijn publiek beschikbaar voor alle gebruikers. Na het verwerken van deze datasets door de oplossing, heeft de onderzoeker de mogelijkheid om de geschatte coördinaten te uploaden. Het benchmarking platform verwerkt deze coördinaten waardoor een duidelijk inzicht in het gedrag van de oplossing kan bekomen worden. Indien de oplossing hardware nodig heeft die niet beschikbaar is in de test omgeving, kan deze direct gekoppeld worden met het benchmarking platform en wordt de nodige hardware in het testbed (al dan niet tijdelijk) geïnstalleerd. Heatmaps, grafieken en statistische analyses worden ter beschikking gesteld als visualisatietools. Tot slot geeft het platform de mogelijkheid om een applicatieprofiel te definiëren (m.a.w. gewichten en eisen koppelen aan metrieken) en scores te berekenen op basis van dat profiel.

Tot slot bevat deze dissertatie richtlijnen die betrekking hebben tot het bepalen hoeveel evaluatiepunten nodig zijn om een objectief resultaat weer te geven. Zoals reeds is aangetoond, kunnen evaluatiepunten een duidelijke invloed hebben op de gepresenteerde prestatie resultaten van een IPS. Om die reden worden richtlijnen voorgesteld die het aantal evaluatiepunten kunnen inschatten op basis van het gewenste betrouwbaarheidsinterval. Dankzij de standaardfout van de gemiddelde waarden is een voorspelling van de standaardafwijking (van de gemiddelde waarden) mogelijk voor elke hoeveelheid aan evaluatiepunten. Deze richtlijnen zijn geverifieerd door meerdere oplossingen die getest zijn in meerdere omgevingen gebruik makend van verschillende technologieën.

Summary

In recent years, location-based information becomes indispensable in multiple application domains. Since Line-of-Sight (LoS) is required for Global Positioning System (GPS)-technologies, it is limited to outdoor environments. The fact wireless technologies (like Wi-Fi, Bluetooth Low Energy (BLE), etc.) become more popular in our surroundings, creates opportunities: Indoor Positioning Systems (IPSs). Typically, an IPS consists of two types of nodes. First, anchor nodes are installed at tactical places inside the building. These will act as the reference points of the system. Next, the person or object that needs to be located, wears or is equipped with a mobile node. The IPS will estimate the coordinates of this mobile node. Nowadays, a wide variety of different technologies and approaches can be differentiated. As such, the number of indoor localization solutions is growing and the awareness that a more thorough way of evaluating and comparing is necessary.

Unfortunately, there is currently no standardized evaluation method for comparing these varied solutions. As a result, every solution has been evaluated under individual, not comparable and not repeatable conditions. Additionally, the selection of evaluation points is often biased: the best accessible evaluation points will be selected for the evaluation. Another issue is the reporting of the obtained performance results: the majority of evaluations focus mainly on point accuracy whist ignoring other crucial metrics such as latency, energy consumption, scalability, cost, simplicity, installation time, etc. Moreover, the accuracy numbers are often calculated using statistical metrics like average or median. This way of evaluation and performance reporting makes it impossible to objectively compare multiple solutions with each other.

This dissertation addresses the above mentioned shortcomings by presenting an innovative benchmarking methodology. A handbook defining multiple metrics and benchmarking scenarios in order to evaluate an IPS in an objective manner. Next, a web-based platform is provided which is aligned with the handbook. It includes multiple tools to support the researcher evaluating its solution. The availability to (i) create & share datasets, (ii) upload the estimated coordinates, (iii) visualize & compare performance results and finally (iv) calculate & rank benchmarking scores based on a self-defined application profile, are an added value for the methodology.

Further, it is important to make a distinction between a “positioning“ and a “tracking” system. Both systems try to estimate the location of a mobile node. However, the latter uses additionally history based information in such a way the complete path can be reconstructed. This history based information is crucial for this type of system to determine the latest position of a mobile node. Since this implies the mobile node needs to communicate continuously to keep the location updated, this dissertation is focussed on indoor *positioning* systems (IPSs).

Initially, this work starts with identifying the issues by optimizing a Time of Arrival (ToA) solution in a harsh industrial environment. Since this environment is an open space containing many metal obstacles, wireless signals suffer from multipath fading due to reflections. Multipath fading is one of the biggest reasons why incorrect measurements (of Received Signal Strength Indicator (RSSI), ToA, etc.) are collected. Advanced calibration and filter techniques can detect those incorrect measurements and eliminate them. Filtering these incorrect measurements improves the solution's performance and strengthens its robustness drastically.

The second part of this dissertation presents comprehensive comparisons and analysis between multiple IPSs in varying indoor environments. Three different solutions are analysed: an IEEE 802.15.4 based on ToA, an IEEE 802.15.4 based on RSSI multilateration and finally an IEEE 802.11 solution based on the fingerprinting principle. These are evaluated in three different industrial environments (each having a different type of walls). All evaluations were executed by the same method. The outcome of this work identified several issues. (i) Multiple inherent trade-offs between different metrics have been noticed, which are typically ignored when reporting only the point accuracy of the solution. More specifically, the results show a very clear trade-off between the collected number of measurements (which are directly translated into energy consumption and latency) and the point-level accuracy. (ii) The accuracy results depend strongly on the characteristics of the environment. (iii) Finally, the choice of evaluation points definitely affects the reported accuracy.

Often, IPSs are evaluated in “optimized” environments without any additional (generated) interference. Which is, again, another way of biasing the reported accuracy. To confirm this, in this dissertation, an analysis is presented of the solutions mentioned above, inside a real-life hospital environment whereby the medical equipment generated realistic interference since this hospital was actively in use. This work confirms the statements identified above. Additionally, it investigates the influence of the anchor nodes. The robustness of the solutions is stress-tested by disabling one or more anchor points.

The major part of this research was defining a benchmarking methodology in order to make comparability between multiple solutions possible. The starting point was the creation of a benchmarking handbook. A handbook which describes procedures for enabling objective evaluation of an IPS. The methodology is aligned with the upcoming ISO/IEC 18305 standard “Test and Evaluation of Localization and Tracking systems”. Moreover, the handbook provides a set of validated and standardized experiment-based benchmarking scenarios focused on the evaluation of Radio-Frequency (RF)-based indoor localization solutions. Accompanied with clear definitions and examples of the performance metrics like point & room accuracy, latency, energy consumption, installation cost, repeatability, interference and environmental sensitivity etc.

Next, a web-based benchmarking platform is presented. This platform consists of multiple tools which support the researcher to obtain more objective performance results of his solution. Every user can create new datasets based on a personal configuration. Accurate Roomba-based robots are used to record these datasets in the w-Lab.t II test facility. In this way, repeatability becomes possible. As such, the robustness of the solution can be evaluated under different conditions. These raw datasets (containing RSSI or ToA measurements) are publicly available for all the other users. After processing these datasets using their solution, the researcher has the possibility to upload the estimated coordinates. The benchmarking platform processes these coordinates whereby a clear insight in the behaviour of the solution can be visualized. If the solution includes custom hardware, a direct connection between the solution and the benchmarking platform can be made and the custom hardware will be deployed in our test facility. Heat maps, graphs and statistical analysis are presented as visualization tools. Finally, the benchmarking platform offers the possibility to define an application profile (giving weights and requirements to metrics in order to obtain a personalized metric score calculation) and retrieve a ranking based on this profile.

Finally, this dissertation proposes a five-step guide to determine the amount of necessary evaluation points. As previously demonstrated in this work, evaluation points can definitely influence the presented performance results of the IPS. Therefore, we present a guide whereby the amount necessary can be estimated based on the desired confidence bounds. Due to the standard error of the mean values, a prediction of standard deviation (of the mean value) can be predicted for each amount of evaluation points. This is verified using multiple solutions, evaluated in multiple test environments using different types of technology.

1 Introduction

"Logic will get you from A to Z.
Imagination will get you everywhere."

–Albert Einstein (1879 - 1955)

This chapter provides the context of the conducted research work, identifies challenges within localization research, summarizes the main contributions and outlines the structure of this dissertation. Finally, it also provides an overview of the publications that were authored during this research period.

1.1 Context

Due to satellite navigation systems and small but powerful pocket computers, also called smartphones, the rising trend of personal location-based services like guidance, tracking or navigation became possible [1]. However, the use of these satellite navigation systems (mainly Global Positioning System (GPS)) is limited to outdoor environments, whereas many commercial applications are envisioned in indoor environments. To remedy this, Indoor Positioning Systems (IPSs) are designed to meet the indoor requirements (also called "Indoor Localization Solutions").

1.1.1 Application domains

Location-awareness in indoor environments becomes indispensable in many varying application domains. Not only personal guidance but also rescue missions, military operations, crowd control, asset tracking in hospitals, guarding detainees, etc. would all be well served by the availability of a reliable positioning solution. An overview of different applications is given in Table 1.1 based on [2].

Table 1.1: Overview of different applications which would benefit from the availability of an indoor localization solution.

User type	Application
Civilian users	Detention facilities Museum / shopping centre guidance Private security guards
Fire fighting	Complex building fires Forest fires Residential and apartment building fires Ship fires Subterranean rescue operations Volcanic eruption detection
Industrial	Entrance admittance Localization in underground mining Production control Warehouse management
Law enforcement	Crowd control Hostage rescue Traffic control
Medical	Localization of elderly Tracking assets Tracking doctors
Military	Building clearing Safe navigation Ship boarding Urban combat

1.1.2 Basic principle of Indoor Localization Solutions

To realize the possibility to locate and track people and assets, a wireless network in this environment will be necessary. As an example, this can be a Wireless Sensor Network (WSN). WSN are networks that contain a huge amount of small and cheap sensor nodes that are communicating through a low-power radio interface. Installing such a network enables many applications. Traditionally, WSNs are used for large-scale sensing tasks and they are relaying the sensed information to a central base station, where the collected information is analysed [3]. E.g. in an underground mine, a sensor network could measure multiple parameters, the temperature, humidity, gas detection, etc. These sensor nodes would be attached at the walls of the tunnel. Additionally, if people and assets are equipped with a sensor node, a location estimation of this sensor node could be determined based on its communication with the fixed sensor nodes at the wall.

This principle of localization is comparable with the GPS. This is illustrated in Figure 1.1. As already mentioned, two types of nodes should be distinguished in the approach:

- **Mobile node(s):** The person or object that needs to be located, wears or is equipped with a mobile node. The solution calculates the location of this mobile device.

- **Anchor node(s):** In order to determine a location of the mobile node, multiple reference points are necessary. For outdoor navigation systems like GPS, satellites are used. Since their orbit is perfectly known, a GPS-receiver can calculate its own location based on the signals received from at least four satellites. For IPSs, the anchor nodes can be the existing Wi-Fi Access Point (AP) in the hospital or additionally installed nodes at tactical locations. Multiple technologies like Bluetooth Low Energy (BLE), Radio-Frequency Identification (RFID), ZigBee, etc. can be used. This will be further elaborated in Section 1.2.

Typically, an IPS consists of an algorithm that processes wireless data from a specific technology. As such, an IPS can be seen as a combination of a localization algorithm running on top of a certain wireless hardware technology. Figure 1.2 presents the different layers.

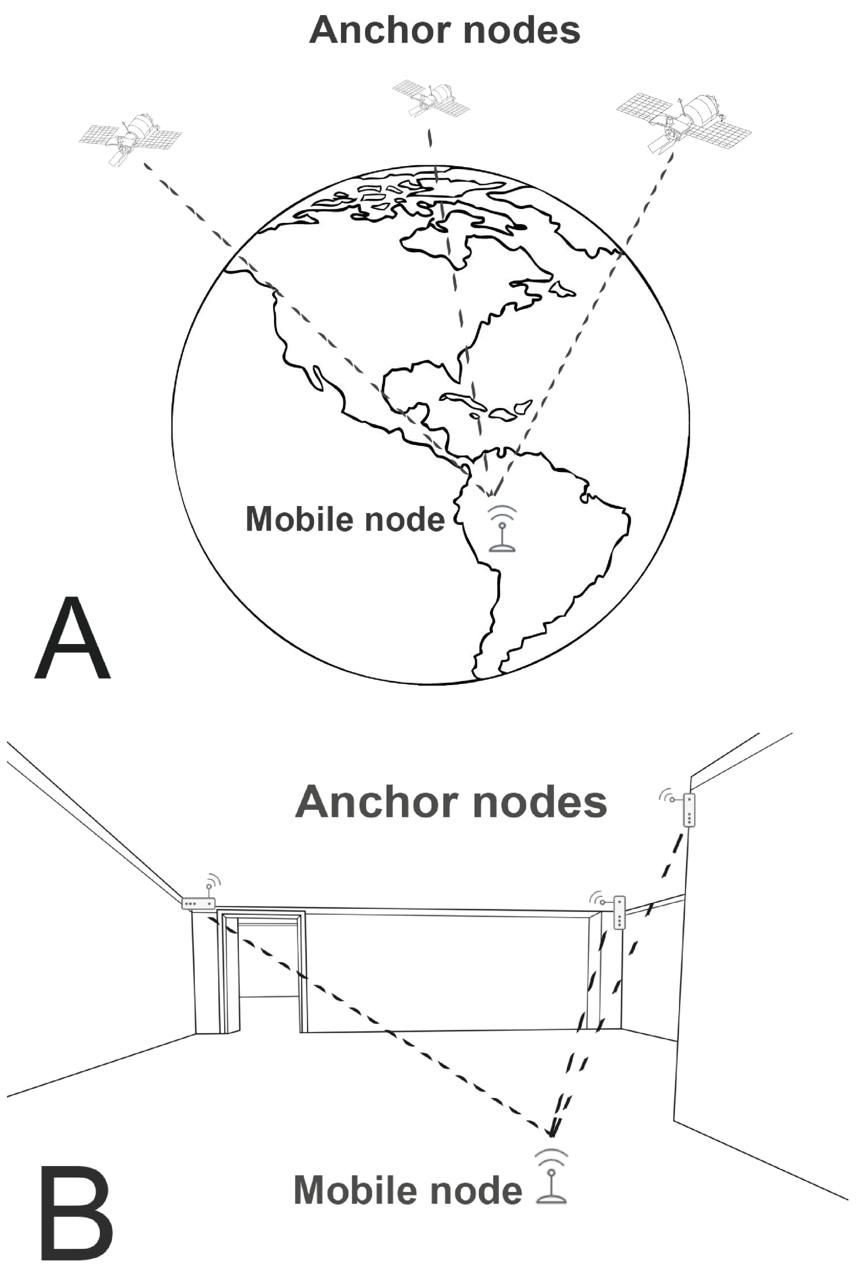


Figure 1.1: (A) The main principle of GPS illustrated, satellites act as anchor nodes broadcasting their signals to the mobile nodes. Mobile nodes receive those signals and estimate their position. In (B) a similar approach is illustrated indoor. Anchor nodes are placed inside a building which communicate with the mobile node.

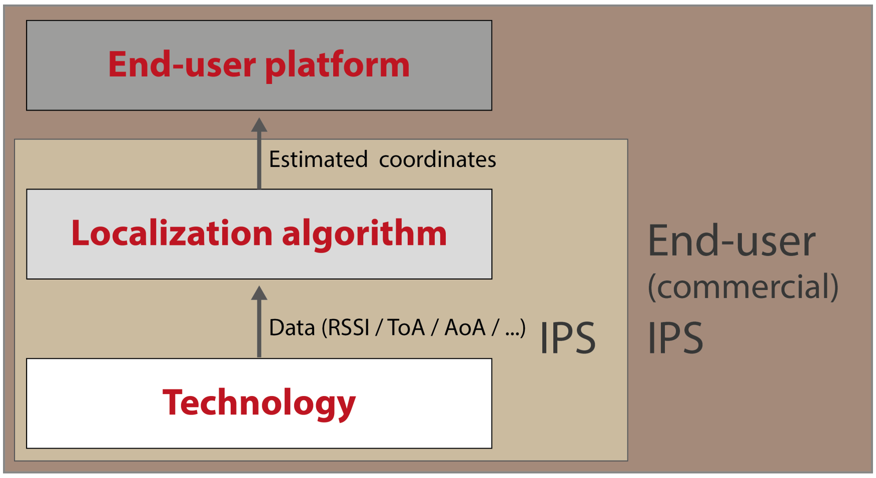


Figure 1.2: The three layers that define an end-user (commercial) IPS. The focus of this PhD are the two lower layers: the technology in combination with a localization algorithm. These two layers define an IPS which delivers estimated coordinates of the mobile node as a result.

1.2 Classification of RF-based Indoor Localization Algorithms

A wide range of approaches are available in order to realize a position estimation inside a building. First, a distinction between an indoor *positioning* system and an indoor *tracking* system can be made. The latter solutions additionally use history based information to estimate the location of the mobile node. As a result, the mobile node needs to communicate continuously to keep the location updated. This has direct implications for the battery lifetime. Further, a start reference point is essential to start tracking. If the starting point is incorrect, the estimated path will be useless. For those reasons, this dissertation is limited to indoor *positioning* systems (IPSs).

In literature [4–11], different approaches can be differentiated: geometric, scene analysis or proximity. Each category will be discussed more in detail in Subsection 1.2.1, 1.2.2 and 1.2.3 respectively. An overview of the classification can be found in Figure 1.3.

1.2.1 Geometric

Geometric (properties of triangles) is the collective name for lateration and angulation techniques. The former estimates the position of the mobile

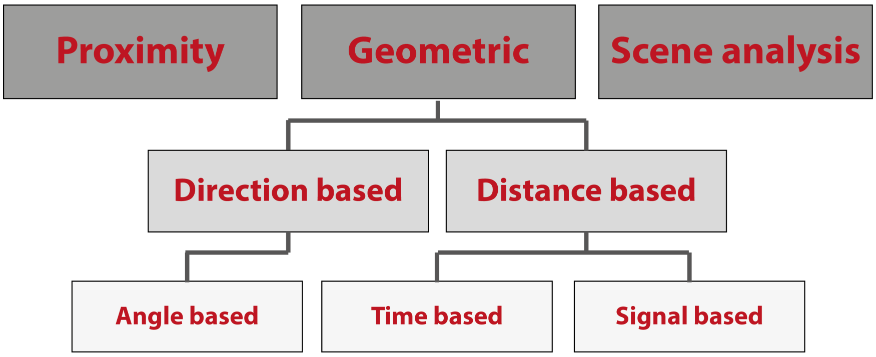


Figure 1.3: Three approaches can be differentiated: proximity, geometric and scene analysis. Further, geometric approaches has two derivations: direction and distance based. Finally, distance based solution can use time or signal based information.

node by measuring its distances from multiple anchor points. The latter computes angles relative to anchor nodes to locate a mobile node. Some research papers define these approaches as range-based ranging techniques [12, 13].

1.2.1.1 Distance based (lateration techniques)

Signal based The main idea of Received Signal Strength Indicator (RSSI) [14] values is that the transmission power P_T directly affect the received power P_R of a signal. Via the Friis transmission equation a linear relationship between both powers in free space becomes clear. As a conclusion, a distance estimation can be derived directly from the obtained RSSI values.

$$P_R = P_T * G_T * G_R \left(\frac{\lambda}{4\pi d}\right)^2 \tag{1.1}$$

where G_T and G_R are the gain of transmitter and receiver respectively. $G_T = G_R = 1$ is usually taken. λ is the wavelength of the signal and d is the distance between sender and receiver. The Path Loss (PL) value can be defined as the ratio of the transmitted power to the received power which is defined as the Free-space path loss (FSPL).

$$PL(dB) = -10\log\left(\frac{P_T}{P_R}\right) = -10\log\left(\frac{\lambda}{4\pi d}\right)^2 = -10\log\left(\frac{1}{FSPL}\right) \tag{1.2}$$

Time based Time of Arrival (ToA), or also called Time of Flight (ToF) [15] estimates distances between devices based on the propagation time of a

Radio-Frequency (RF) wave from one transmitter to one receiver. With the speed of light, the distance can be calculated. One or two way ToA is possible. One way requires a very accurate clock synchronisation between transmitter and receiver. This is rather hard to achieve with low-end devices. The solution is to apply two way ToA. In this situation the single Round Trip Time (RTT) is measured and clock synchronization is no longer necessary. A variant to this is the Symmetrical Double-Sided Two-Way Ranging (SDS-TWR) [16]. This variant further improves the accuracy by sending one additional message per ranging measurement. However it assumes the anchor nodes also perform and collect ranging measurements. If this is the case, additional communication between sender and receiver is unavoidable.

Basically, two way ToA is applied. A distance estimation can be derived based on

$$\begin{aligned} d &= \frac{c}{2} t_{flight} \\ &= \frac{c}{2} \left(t_{RTT} - t_{delay} \right) \end{aligned} \tag{1.3}$$

Hereby, c is the speed of light, t_{flight} the measured time it takes for the signal to propagate, t_{RTT} the calculated RTT and t_{delay} including all the time necessary for sending the packet, processing the packet and generating the Acknowledgement (ACK).

TDoA Time-Difference of Arrival (TDoA) [17, 18] its purpose is to determine the relative position of the mobile node (in this case the transmitter) by examining the time difference of the arrival of the signal at multiple anchor nodes. The anchors must be synchronized accurately since they receive the signal each at a different time. It is this difference which is used to calculated the position of the mobile transmitter. As an 2D example, assume mobile node X transmits a message, which arrives at receiver A with time T_A and at receiver B at time T_B. The time difference of this message at two receivers A and B will be calculated as a constant k_1. This will narrow the possible position of the mobile node to a hyperbola, denoted as $TDOA_{B-A}$.

$$TDOA_{B-A} = |T_B - T_A| = k_1 \tag{1.4}$$

This hyperbola represents all the possible locations of the mobile transmitter based on those two receivers. The knowledge of a difference of time of arrival calculated using a third receiver will create hyperbola $TDOA_{C-A}$. The intersection of those hyperbolas is the estimated position of the mobile node. Figure 1.4 depicts this situation.

$$TDOA_{C-A} = |T_C - T_A| = k_2 \tag{1.5}$$

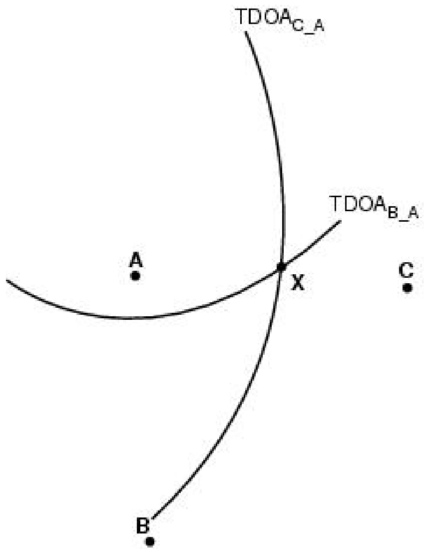


Figure 1.4: Basic scheme for TDoA, mobile node X transmits a message, which is received by receivers A, B and C. The two hyperbola's show the constant time difference value. The intersection of those hyperbola's is the estimated location of mobile node X.

1.2.1.2 Angle based (angulation technique)

AoA With Angle of Arrival (AoA) [19, 20], the location of the mobile transmitter can be found by the intersection of several pairs of angle direction lines as demonstrated in Figure 1.5. Hereby, antenna arrays are required since the time difference of arrival on each individual antenna element is used to determine the angle.

1.2.2 Scene Analysis

In contrast to the previous described approaches, scene analysis solutions [21, 22] collect features (fingerprints) of a scene and then estimate the location by matching a new measurement with the closest possible match in the database of fingerprints. The core functionality of these algorithms is trying to find the best match between a live measurement and a database of stored measurements. The process is twofold, first, in the offline phase

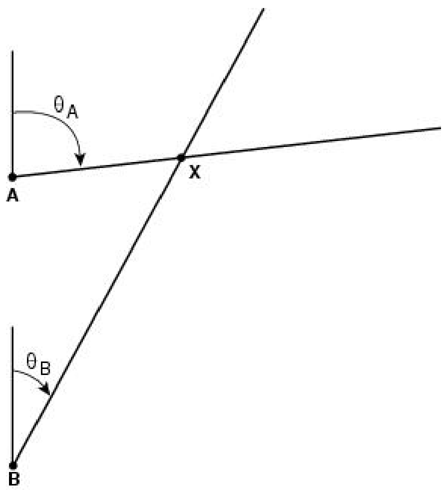


Figure 1.5: At least two anchor nodes are necessary to estimate the location of the mobile transmitter. Based on angles θ_A and θ_B, the location of mobile transmitter X can be determined. However exceptional antennas such as antenna arrays are required at receivers A and B.

(also called training phase) a time consuming site survey in an environment is executed. The location coordinates in combination with information (RSSI,Service Set Identifier (SSID)) about the nearby base stations is stored into a large database. Next, during the online phase, the solution uses the currently observed signals from the environment and compares it with all the fingerprints in the database. The best matching fingerprint its coordinates will be used as the estimated position of the mobile transmitter. Two important approaches will be discussed below.

Probabilistic methods These algorithms consider positioning as a classification problem [23]. Assuming there are n location candidates L_1, L_2, L_3, ..., L_n and s is the observed fingerprint in the online phase. A general decision can be determined (Equation 1.6).

$$\text{Choose } L_i \text{ if } P(L_i|s) > P(L_j|s) \text{ for } i, j = 1, 2, 3, ..., n \text{ with } j \neq i \qquad (1.6)$$

Here, $P(L_i|s)$ denotes the probability that the mobile node is at location L_i given that the observed fingerprint is s. $P(L_i)$ is the probability that the mobile node is at location L_i. The decision rule above is based on posteriori probability. Based on Bayes' formula and assuming that $P(L_i) = P(L_j)$ for $i, j = 1, 2, 3, ..., n$ the following decision rule can be formulated

(Equation 1.7). Here, $P(s|L_i)$ is the probability that the signal vector s is received, given that the mobile node is located in location L_i.

$$\text{Choose } L_i \text{ if } P(s|L_i) > P(s|L_j) \text{ for } i, j = 1, 2, 3, ..., n \text{ with } j \neq i \qquad (1.7)$$

However, this is applicable for discrete location candidates (which are stored in the fingerprinting database). The final estimated 2D location $(\hat{x}, \hat{y})$ can be calculated using interpolation: a weighted average of the coordinates (x_L, y_L) of all sampling locations (Equation 1.8).

$$(\hat{x}, \hat{y}) = \sum_{i=1}^{n} (P(L_i|s)(x_L, y_L)) \qquad (1.8)$$

kNN The k-Nearest Neighbours (kNN) principle [24] uses an online fingerprint s to search for the k closest matches from the database built during the offline phase according to the root mean square errors principle. This principle is a classification algorithm whereby the k-closest selected fingerprints are averaged in order to retrieve the estimated coordinates. A visualization of this classification can be found in Figure 1.6.

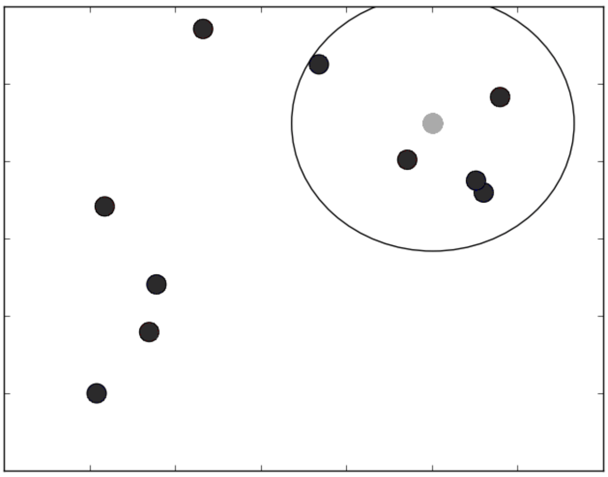


Figure 1.6: The online sample s is marked as the light gray on the map. Black dots represent fingerprints which are stored in the database gathered during the offline phase. Based on the root mean square error principle, the k-closest neighbours are selected and averaged to estimate the location of the mobile node. In this case, $k = 5$ and thus the circle collects the five closest neighbours.

1.2.3 Proximity

Proximity based solutions [25, 26] use a weak sending power. If a message is received, the receivers knows it is in the vicinity of the sender. A rough estimation is using the senders' coordinates. For many applications, this can be sufficient. In this way, a simple, easy to implement, solution can be provided. Proximity solutions are in combination with the scene analysis approaches in literature also known as the range-free solutions [27] since no distance estimation is calculated to estimate the mobile node's position. A well known variant of proximity, is the centroid solution.

Centroid A centroid algorithm [28] can be applied when a mobile node of a proximity solution receives messages from multiple receivers simultaneously. The estimated coordinates $(\hat{x}, \hat{y})$ of the mobile node is the centroid (x_M, y_M) of all the anchor nodes which received a beacon.

$$(\hat{x}, \hat{y}) = (x_M, y_M) = \left(\frac{\sum_{n=0}^{k} x_n}{k} \; , \; \frac{\sum_{n=0}^{k} y_n}{k} \right) \tag{1.9}$$

1.2.4 Additional categories

Previous subsections classified indoor localization approaches which make use of RF-based technologies. However since the computational power of smartphones is increasing year by year, more solutions are focussing on the possibilities of a smartphone [29, 30]. Besides the RF-technologies in smartphones, also the camera [31–33] and the Inertial Measurement Unit (IMU) [34, 35] can be used to estimate the position of the smartphone and thus also the person who is locating himself.

1.2.5 Indoor Positioning Solutions categorized by technology

This section provides a brief overview of common IPSs categorized by technology. The first two technologies are not RF-based.

Infra-Red (IR) IR localization [36, 37] uses beacons which are transmitted in the infra-red spectrum. As such, these signals cannot travel through walls, causing low coverage. Moreover this technology is sensitive to sunlight. Examples are (i) Active Badge [38], designed at AT&T Cambridge. An active badge transmits a globally unique IR signal every 15 seconds. Each room contains one or more sensors to pick up this IR signal. (ii) Firefly [39], which is designed by Cybernet System Corporation, is an IR-based motion tracking system. Tags emit IR-light which is collected by three camera's which act as a camera array. It can offer an accuracy of about 3.0mm. (iii) OPTOTRAK PRO [40] is focussed on congested shops and workspaces and is based on the same principle as Firefly. Tags emit IR light that is detected by three camera's placed on an array.

Ultrasonic People were inspired by bats, which use ultrasound signals to navigate in the night. Ultrasonic technology can be used in geometric based solutions [41]. However, these signals cannot go through walls, like IR. A first example is (i) Active Bat [42, 43], also designed by AT&T Cambridge. It uses ultrasound in combination with a triangulation location technique. The tag periodically broadcasts a short pulse of ultrasound, which is received by a matrix of ceiling mounted receivers. Due to the ToA, the distance between sender and receivers can be calculated. (ii) Cricket [44] is an often used indoor localization example. The cricket system uses both ultrasonic and RF signals to calculate the distances. The purpose of the RF messages is synchronisation of the ToA measurements and forwarding their location information in a decentralized way. Other examples are DOLPHIN [45, 46] and Sonitor [47].

RFID RFID positioning [48–50] is often used in complex indoor environments such as offices and hospitals. Commonly in combination with proximity whereby the user has to identify himself with an RFID-tag. A distinction between passive and active RFID can be made. With passive RFID, the tag to be tracked is the receiver. Active RFID tags are transceivers, which can actively transmit their identification. The following examples are based on the latter group. (i) WhereNet [51], designed by Zebra Technology Company and uses Differential Time of Arrival (DToA) to estimate the location of their tags. Other examples are (ii) SpotON [52] and (iii) LANDMARC [53].

Wi-Fi A very popular and grateful technology is Wi-Fi since it has already been implemented in almost every environment. Wi-Fi based positioning [54–57] can reuse the existing network which is a great advantage. Mostly, RSSI values of the signals are used to determine the mobile node's location like (i) RADAR [58], (ii) COMPASS [59] and (iii) Ekahau [60]. Altough some solutions apply a ToA approach [61]. Finally, fingerprinting solutions [62] eagerly use Wi-Fi networks like Redpin [63].

BLE Bluetooth is the new rising trend as chosen technology for indoor positioning [64–66]. Thank to the advertising mode, BLE beacons can very easily broadcast short messages at flexible update rates. Commercial examples are (i) Indoo.rs [67], (ii) Blooloc [68] and InfSoft [69].

WSN Sensor networks are often used for RSSI based indoor positioning. Unlike Wi-Fi, the RSSI is specified in the IEEE 802.15.4 standard. This is the most widely used radio on today's sensor node platforms. Many solutions in literature use WSN [70–72].

Ultra-Wideband (UWB) Finally, UWB gains popularity [73–78] since the cost of an UWB-radio drops. With UWB technology, high accuracy positioning becomes possible. It spreads information over a large spectrum in pulses. As such, it is insensitive for interference. A commercial example of indoor positioning with UWB is Pozyx [79].

1.3 Problem statement

The previous section gives the current state of the art of indoor localization solutions and proofs the amount of published research has increased tremendously over the last few years. Unfortunately, there is currently no standardized evaluation method for comparing these varied solutions. As a result, each researcher evaluates their own solution in their own, proprietary environment. In other words, the current problem is the lack of comparability between solutions due to the fact that they have been evaluated under individual, not comparable and not repeatable conditions.

Further, due to the complexity of evaluating indoor localization solutions at scale, many solutions remain either theoretical or are evaluated in easily-accessible environments without any (additional) RF interference. Selecting easily-accessible evaluation points can definitely bias the presented performance results.

In addition, the majority of evaluations focus mainly on point accuracy of the result whilst ignoring crucial application-level metrics such as delay, energy consumption, scalability, cost, simplicity, installation time, etc. Moreover, the reported accuracy is often calculated using different calculation statistics (average, median, percentiles, etc.) making comparing almost impossible.

These shortcomings were also stated and addressed in the Evaluation of RF-based Indoor localization Solutions (EVARILOS)-project [80]: the pitfall to reproduce research results in real life scenarios suffering from uncontrolled RF interference and the weakness of numerous published solutions being evaluated under not comparable conditions. The major outcomes of this project were a benchmarking handbook [81], a benchmarking platform [82] and an open challenge [83].

1.3.1 Relevant performance metrics

Currently, only point accuracy results are presented as the performance of an indoor localization solution. Though for a comprehensive comparison, other relevant metrics have to be taken into account. Therefore, this section gives an overview of the most important metrics beside point accuracy.

- **Room accuracy**: the possibility to locate a stationary mobile node at room level. For many applications, this degree of accuracy is already acceptable. E.g. locating an important but rarely used medical device can be equipped with a mobile node. If the position of this device is requested, room (and thus also floor) information can be sufficient.

- **Latency**: the time between sending a location request and receiving the location information. To continue with the previous example: the time it takes starting when a staff member sends a location request to locate a mobile node (which can be carried with a patient or attached to a medical device) until the staff member receives the location information that was requested. Another example is the "emergency call". When a patient pushed a mobile panic button, the latency of the localization solution can have an impact on the health status of the involved patient.

- **Installation time / cost**: mostly, environments are (almost) continuously operational, meaning the installation time must be reduced to the minimum. Can the existing network be reused or is new wiring necessary? Does the solution requires recalibrating or not? Answers to those questions are reflected in the installation time and cost metric.

- **Energy consumption**: this metric is particularly important for the mobile node. This value is equivalent with the life-time of the device. Most of the time, this metric is directly proportional with the latency / update cycle metric.

- **Robustness**: a twofold metric: on the one hand, the resilience of the system. How robust is the system if anchor nodes fail? Does adding an anchor node into system requires a complete recalibration? On the other hand, a special focus on interference robustness is crucial. Mostly, solutions are evaluated in specially designed test environments which have (almost) perfect circumstances. “In the field”, these environments are far from perfect. Many other technologies are already operational, people are walking around, assets are being moved all the time. This has to be taken into account.

- **Low cost**: every commercial application has to keep an eye on the total price of the system. The added value of the system must dominate the price tag. The necessary number of anchor nodes, the chosen technology, installation time, calibration requirements, etc. all influence this metric.

- **Scalability**: this metric includes the ability of the system, network or process to handle growing amount of work in an efficient and capable manner. It should manage expanding the environment or the amount of mobile nodes which should be localized simultaneously.

1.4 Outline of PhD dissertation

This dissertation is composed of a number of publications that were realized within the scope of this PhD. The selected publications provide an integral and consistent overview of the work performed. The different research contributions are detailed in Section 1.5 and the complete list of publications that resulted from this work is presented in Section 1.6.

Within this section we give an overview of the remainder of this dissertation and explain how the different chapters are linked together. Figure 1.7 positions the different contributions that are presented in each chapter (CH).

As visualized, this dissertation can be divided in three different parts. In order to start benchmarking and comparing multiple indoor localization solutions, a good insight in this topic is required. Therefore, the first part, Part I (Chapter 2) handles about a contribution in optimizing the performance of a ToA indoor localization solution in harsh industrial environments. The ranging measurements of ToA can be influenced by its environment. Certainly when a lot of multi-path fading causes reflections in the environment. Advanced filter techniques can reduce the error distances drastically. The entire process to optimize these issues is elaborated in this chapter.

The second part, Part II, includes two chapters which try to compare multiple solutions in an objective way. Both are experimental studies. Chapter 3 is focussed on heterogeneous industrial environments (different wall

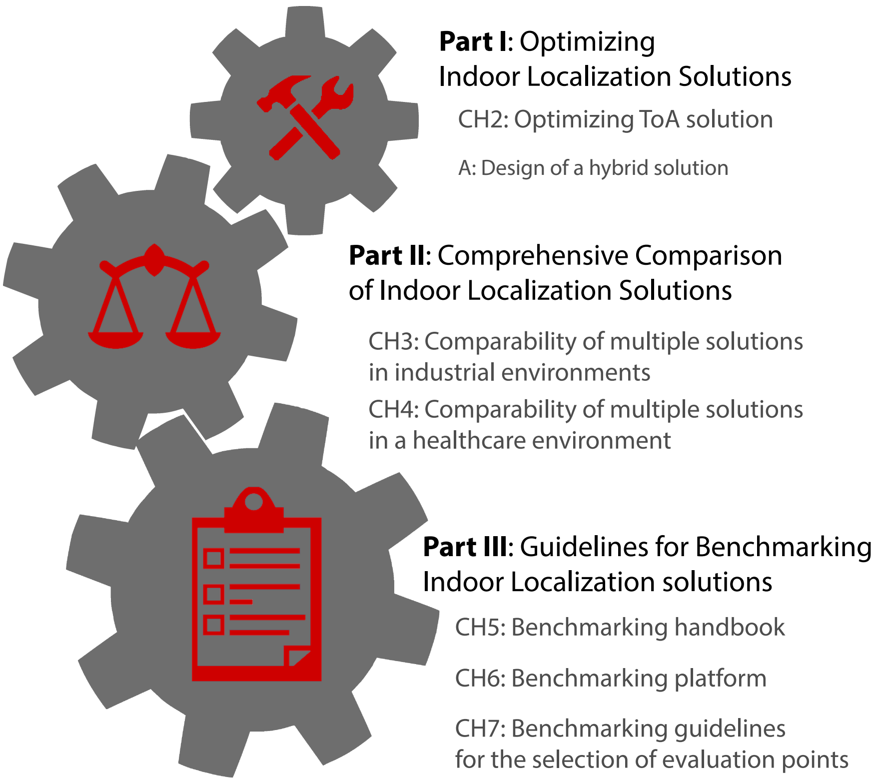


Figure 1.7: Schematic position of the different chapters in this dissertation

types) whilst Chapter 4 especially is focussed on performance analysis in a healthcare environment. Both chapters start from the initial version of the handbook and platform elaborated in Part III.

Based on the outcome of Part II, the shortcomings of these tools became clear. Part III bundles the available tools. Chapter 5 gives the researcher a better insight in how to benchmark his solution. The definitions, methodology, scenarios and workflows are elaborated in this chapter. In order to support the researcher as much as possible, additional software tools are provided. These tools are gathered in a benchmarking platform. The functionality and demonstration of this platform is written in Chapter 6. Still, some parts in the benchmarking methodology were lacking. As an example, the essential amount of evaluation points in order to represent a reliable performance result, was insufficiently described. Chapter 7 investigates this subject and delivers a five-step guide to achieve a sufficient amount of evaluation points.

Finally, overall conclusions of this PhD can be found in Chapter 8.

1.5 Research contributions

In Section 1.3, the problems and challenges for benchmarking and comparing multiple indoor localization solutions are formulated. They are tackled in the remainder of this PhD dissertation for which the outline is given in Section 1.4. To conclude, we present an elaborated list of the research contributions within this dissertation:

- **Contribution 1**: Optimizing existing indoor localization solutions
  This first contribution was indispensable to get familiar with the field of study. Existing solutions are analysed thoroughly and optimized applying filter techniques. These solutions are comprehensive evaluated in our wilab-environments.

- **Contribution 2**: Defining a benchmarking methodology
  With the lessons learned from the first contribution, a benchmarking methodology was designed which defined definitions, guidelines, scenarios and workflows in order to benchmark an indoor localization solution as objective as possible.

- **Contribution 3**: Developing a benchmarking platform
  To support the benchmarking methodology, several tools were developed which researchers can use. It is a web-based platform which enables an automated evaluation and comparison of multiple solutions in different environments and using multiple evaluation metrics.

The platform and the plugins can be used in real-time on existing wireless testbed facilities, while also supporting a remote offline evaluation method using precollected data traces. Finally, it provides intuitive visualization and comparing tools of the obtained results.

- **Contribution 4**: Designing a guide for determining the amount of evaluation points
  A final contribution is the statistical analyses to determine the necessary amount of evaluation points in order to represent objective reliable results. Based on the standard error of the mean value, a five-step guide could be designed.

1.6 Publications

The research results obtained during this PhD research have been published in scientific journals and presented at a series of international conferences. The following list provides an overview of the publications during my PhD research.

1.6.1 Publications in international journals (listed in the Science Citation Index [1])

1. **Tom Van Haute**, Eli De Poorter, Ingrid Moerman, Filip Lemic, Vlado Handziski, Adam Wolisz, Niklas Wirström, Thiemo Voigt. *Comparability of RF-based Indoor Localization Solutions in Heterogeneous Environments: An Experimental Study.* Published in the International Journal of Ad Hoc and Ubiquitous Computing, (in press), April 2015.
2. **Tom Van Haute**, Eli De Poorter, Filip Lemic, Vlado Handziski, Niklas Wirström, Thiemo Voigt, Adam Wolisz, Ingrid Moerman. *Platform for Benchmarking of RF-Based Indoor Localization Solutions*. Published in IEEE Communications Magazine: Network Testing, 53.9: 126-133, September 2015.
3. **Tom Van Haute**, Eli De Poorter, Pieter Crombez, Filip Lemic, Vlado Handziski, Niklas Wirström, Adam Wolisz, Thiemo Voigt, Ingrid Moerman. *Performance Analysis of Multiple Indoor Positioning*

[1] The publications listed are recognized as 'A1 publications', according to the following definition used by Ghent University: A1 publications are articles listed in the Science Citation Index Expanded, the Social Science Citation Index or the Arts and Humanities Citation Index of the ISI Web of Science, restricted to contributions listed as article, review, letter, note or proceedings paper.

Systems in a Healthcare Environment. Published in the International Journal of Health Geographics, 15.1: 1, February 2016.

4. **Tom Van Haute**, Bart Verbeke, Eli De Poorter, Ingrid Moerman. *Optimizing Time of Arrival Localization Solutions for Challenging Industrial Environments*. Published in IEEE Transactions on Industrial Informatics, (in press), April 2016.

5. **Tom Van Haute**, Eli De Poorter, Eric Laermans, Ingrid Moerman. *Benchmarking of Localization Solutions: Guidelines for the Selection of Evaluation Points* Submitted to Elsevier Ad Hoc Networks, May 2016.

1.6.2 Publications in international conferences

1. **Tom Van Haute**, Eli De Poorter, Jen Rossey, Ingrid Moerman, Vlado Handziski, Arash Behboodi, Filip Lemic, Adam Wolisz, Niklas Wiström, Thiemo Voigt, Pieter Crombez, Piet Verhoeve, Jose Javier de las Heras. *The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions.* Published in the proceedings of the MERMAT workshop, Ireland, Dublin, 2013.

2. Arash Behboodi, Pieter Crombez, Jose Javier de las Heras, Eli De Poorter, Vlado Handziski, Filip Lemic, Ingrid Moerman, **Tom Van Haute**, Piet Verhoeve, Thiemo Voigt, Niklas Wirström, Adam Wolisz. *Poster Abstract: Evaluation of RF-based Indoor Localization Solutions for the Future Internet.* Published in the proceedings of the Future Network and Mobile Summit, Portugal, Lisbon, 2013.

3. **Tom Van Haute**, Jen Rossey, Pieter Becue, Eli De Poorter, Ingrid Moerman. *A hybrid indoor localization solution using a generic architectural framework for sparse distributed wireless sensor networks.* Published in the proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS), Warsaw, Poland, 2014.

4. Filip Lemic, Vlado Handziski, Niklas Wirström, **Tom Van Haute**, Eli De Poorter, Thiemo Voigt, Adam Wolisz. *Demo Abstract: Virtual Experimental Evaluation of RF-based Indoor Localization Algorithms.* Published in the proceedings of the 12th European Conference on Wireless Sensor Networks (EWSN'15), Porto, Portugal, 2015.

5. Filip Lemic, Vlado Handziski, Niklas Wirström, **Tom Van Haute**, Eli De Poorter, Thiemo Voigt, Adam Wolisz. *Web-based Platform for*

Evaluation of RF-based Indoor Localization Algorithms. Published in the proceedings of the IEEE ICC 2015 Workshop on Advances in Network Localization and Navigation (ANLN), London, UK, 2015.

6. Filip Lemic, Vlado Handziski, Adam Wolisz, Giuseppe Caso, Luca De Nardis, Pieter Crombez, **Tom Van Haute**, Eli De Poorter. *Toward Extrapolation of Wi-Fi Fingerprinting Performance Across Environments.* Published in the proceedings of the 17th Workshop on Mobile Computing Systems and Applications (ACM HotMobile'16), St. Augustine, FL, USA, 2016.

7. **Tom Van Haute**, Eli De Poorter, Filip Lemic, Vlado Handziski, Niklas Wirstrm̈, Adam Wolisz, Ingrid Moerman. *Demonstration Abstract: Platform for Benchmarking RF-Based Indoor Localization Solutions.* Published in the proceedings of the 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN'16), Vienna, Austria, 2016.

References

[1] N. M. Drawil, H. M. Amar, O. Basir, et al. *GPS localization accuracy classification: A context-based approach*. Intelligent Transportation Systems, IEEE Transactions on, 14(1):262–273, 2013.

[2] J. Rantakokko, P. Händel, M. Fredholm, and F. Marsten-Eklöf. *User requirements for localization and tracking technology: a survey of mission-specific needs and constraints*. In Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference On, pages 1–9. IEEE, 2010.

[3] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury. *A survey on wireless multimedia sensor networks*. Computer networks, 51(4):921–960, 2007.

[4] H. Koyuncu and S. H. Yang. *A survey of indoor positioning and object locating systems*. IJCSNS International Journal of Computer Science and Network Security, 10(5):121–128, 2010.

[5] Y. Gu, A. Lo, and I. Niemegeers. *A survey of indoor positioning systems for wireless personal networks*. Communications Surveys & Tutorials, IEEE, 11(1):13–32, 2009.

[6] H. Liu, H. Darabi, P. Banerjee, and J. Liu. *Survey of wireless indoor positioning techniques and systems*. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37(6):1067–1080, 2007.

[7] R. Mautz. *Overview of current indoor positioning systems*. Geodezija ir kartografija, 35(1):18–22, 2009.

[8] S. Pandey and P. Agrawal. *A survey on localization techniques for wireless networks*. Journal of the Chinese Institute of Engineers, 29(7):1125–1148, 2006.

[9] S. Adler, S. Schmitt, K. Wolter, and M. Kyas. *A survey of experimental evaluation in indoor localization research*. In Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on, pages 1–10. IEEE, 2015.

[10] D. Zhang, F. Xia, Z. Yang, L. Yao, and W. Zhao. *Localization technologies for indoor human tracking*. In Future Information Technology (FutureTech), 2010 5th International Conference on, pages 1–6. IEEE, 2010.

[11] N. R. Mrindoko and L. M. Minga. *A Comparison Review of Indoor Positioning Techniques*. International Journal of Computer (IJC), 21(1):42–49, 2016.

[12] X. Luo, W. J. OBrien, and C. L. Julien. *Comparative evaluation of Received Signal-Strength Index (RSSI) based indoor localization techniques for construction jobsites*. Advanced Engineering Informatics, 25(2):355–363, 2011.

[13] Y. Zhou, C. L. Law, Y. L. Guan, and F. Chin. *Indoor elliptical localization based on asynchronous UWB range measurement*. Instrumentation and Measurement, IEEE Transactions on, 60(1):248–257, 2011.

[14] S. Mazuelas, A. Bahillo, R. M. Lorenzo, P. Fernandez, F. A. Lago, E. Garcia, J. Blas, and E. J. Abril. *Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks*. Selected Topics in Signal Processing, IEEE Journal of, 3(5):821–831, 2009.

[15] M. Ciurana, S. Cugno, and F. Barceló-Arroyo. *WLAN indoor positioning based on TOA with two reference points*. In Positioning, Navigation and Communication, 2007. WPNC'07. 4th Workshop on, pages 23–28. IEEE, 2007.

[16] Y. Jiang and V. Leung. *An asymmetric double sided two-way ranging for crystal offset*. In Signals, Systems and Electronics, 2007. ISSSE'07. International Symposium on, pages 525–528. IEEE, 2007.

[17] F. Gustafsson and F. Gunnarsson. *Positioning using time-difference of arrival measurements*. In Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on, volume 6, pages VI–553. IEEE, 2003.

[18] M. Bocquet, C. Loyez, and A. Benlarbi-Delai. *Using enhanced-TDOA measurement for indoor positioning*. Microwave and Wireless Components Letters, IEEE, 15(10):612–614, 2005.

[19] R. Peng and M. L. Sichitiu. *Angle of arrival localization for wireless sensor networks*. In Sensor and Ad Hoc Communications and Networks, 2006. SECON'06. 2006 3rd Annual IEEE Communications Society on, volume 1, pages 374–382. IEEE, 2006.

[20] P. Kułakowski, J. Vales-Alonso, E. Egea-López, W. Ludwin, and J. García-Haro. *Angle-of-arrival localization based on antenna arrays for wireless sensor networks*. Computers & Electrical Engineering, 36(6):1181–1186, 2010.

[21] J. Torres-Sospedra, R. Montoliu, S. Trilles, Ó. Belmonte, and J. Huerta. *Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems*. Expert Systems with Applications, 42(23):9263–9278, 2015.

[22] G. Caso and L. De Nardis. *On the applicability of Multi-Wall Multi-Floor propagation models to WiFi Fingerprinting Indoor Positioning*. In Future Access Enablers for Ubiquitous and Intelligent Infrastructures, pages 166–172. Springer, 2015.

[23] P. Kontkanen, P. Myllymäki, T. Roos, H. Tirri, K. Valtonen, and H. Wettig. *Topics in probabilistic location estimation in wireless networks.* In PIMRC, pages 1052–1056, 2004.

[24] J. Ma, X. Li, X. Tao, and J. Lu. *Cluster filtered KNN: A WLAN-based indoor positioning scheme*. In World of Wireless, Mobile and Multimedia Networks, 2008. WoWMoM 2008. 2008 International Symposium on a, pages 1–8. IEEE, 2008.

[25] P. Deepesh, R. Rath, A. Tiwary, V. N. Rao, and N. Kanakalata. *Experiences with using iBeacons for Indoor Positioning*. In Proceedings of the 9th India Software Engineering Conference, pages 184–189. ACM, 2016.

[26] C.-H. Huang, L.-H. Lee, C. C. Ho, L.-L. Wu, and Z.-H. Lai. *Real-time RFID indoor positioning system based on Kalman-filter drift removal and Heron-bilateration location estimation*. Instrumentation and Measurement, IEEE Transactions on, 64(3):728–739, 2015.

[27] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. *Range-free localization schemes for large scale sensor networks*. In Proceedings of the 9th annual international conference on Mobile computing and networking, pages 81–95. ACM, 2003.

[28] G.-A. L. Zodi, G. P. Hancke, G. P. Hancke, and A. B. Bagula. *Enhanced centroid localization of wireless sensor nodes using linear and neighbor weighting mechanisms*. In Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication, page 43. ACM, 2015.

[29] M. Weyn, M. Klepal, et al. *Adaptive motion model for a smart phone based opportunistic localization system*. In Mobile Entity Localization and Tracking in GPS-less Environnments, pages 50–65. Springer, 2009.

[30] M. Klepal, M. Weyn, W. Najib, I. Bylemans, S. Wibowo, W. Widyawan, and B. Hantono. *Ols: opportunistic localization system for smart phones devices*. In Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds, pages 79–80. ACM, 2009.

[31] J. Choi and G. Friedland. *Multimodal Location Estimation of Videos and Images*. Springer, 2015.

[32] A. Mulloni, D. Wagner, I. Barakonyi, and D. Schmalstieg. *Indoor positioning and navigation with camera phones*. Pervasive Computing, IEEE, 8(2):22–31, 2009.

[33] H. Hile and G. Borriello. *Positioning and Orientation in Indoor Environments Using Camera Phones.* IEEE Computer Graphics and Applications, 28(4):32–39, 2008.

[34] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao. *A reliable and accurate indoor localization method using phone inertial sensors*. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pages 421–430. ACM, 2012.

[35] E. Foxlin. *Pedestrian tracking with shoe-mounted inertial sensors*. Computer Graphics and Applications, IEEE, 25(6):38–46, 2005.

[36] P.-W. Chen, K.-S. Ou, and K.-S. Chen. *IR indoor localization and wireless transmission for motion control in smart building applications based on Wiimote technology*. In SICE Annual Conference 2010, Proceedings of, pages 1781–1785. IEEE, 2010.

[37] E. Aitenbichler and M. Mühlhäuser. *An IR local positioning system for smart items and devices*. In Distributed Computing Systems Workshops, 2003. Proceedings. 23rd International Conference on, pages 334–339. IEEE, 2003.

[38] R. Want, A. Hopper, V. Falcao, and J. Gibbons. *The active badge location system*. ACM Transactions on Information Systems (TOIS), 10(1):91–102, 1992.

[39] Cybernet Systems. *Firefly motion tracking*. http://www.cybernet.com/interactive/firefly/, 2016. [Online; accessed 05-2016].

[40] NDI. *Firefly motion tracking*. http://www.ndigital.com/msci/products/procmm/, 2016. [Online; accessed 05-2016].

[41] M. Hazas and A. Hopper. *Broadband ultrasonic location systems for improved indoor positioning*. Mobile Computing, IEEE Transactions on, 5(5):536–547, 2006.

[42] A. Ward, A. Jones, and A. Hopper. *A new location technique for the active office*. Personal Communications, IEEE, 4(5):42–47, 1997.

[43] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper. *Implementing a sentient computing system*. Computer, 34(8):50–56, 2001.

[44] N. B. Priyantha. *The cricket indoor location system*. PhD thesis, Massachusetts Institute of Technology, 2005.

[45] Y. Fukuju, M. Minami, H. Morikawa, and T. Aoyama. *DOLPHIN: An Autonomous Indoor Positioning System in Ubiquitous Computing Environment.* In WSTFEUS, pages 53–56, 2003.

[46] M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama. *DOLPHIN: a practical approach for implementing a fully distributed indoor ultrasonic positioning system*. In UbiComp 2004: Ubiquitous Computing, pages 347–365. Springer, 2004.

[47] *Sonitor*. http://www.sonitor.com/, 2016. [Online; accessed 05-2016].

[48] A. Bekkali, H. Sanson, and M. Matsumoto. *RFID indoor positioning based on probabilistic RFID map and Kalman filtering*. In Wireless and Mobile Computing, Networking and Communications, 2007. WiMOB 2007. Third IEEE International Conference on, pages 21–21. IEEE, 2007.

[49] S. S. Saab and Z. S. Nakad. *A standalone RFID indoor positioning system using passive tags*. Industrial Electronics, IEEE Transactions on, 58(5):1961–1970, 2011.

[50] T. Sanpechuda and L. Kovavisaruch. *A review of RFID localization: Applications and techniques*. In Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on, volume 2, pages 769–772. IEEE, 2008.

[51] Zebra. *WhereNet*. https://www.zebra.com/us/en/solutions/location-solutions/enabling-technologies/wherenet.html/, 2016. [Online; accessed 05-2016].

[52] J. Hightower, R. Want, and G. Borriello. *SpotON: An indoor 3D location sensing technology based on RF signal strength*. UW CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA, 1, 2000.

[53] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. *LANDMARC: indoor location sensing using active RFID*. Wireless networks, 10(6):701–710, 2004.

[54] M. Weyn and F. Schrooyen. *A wi-fi assisted gps positioning concept*. ECUMICT, Gent, Belgium, 2008.

[55] S. Woo, S. Jeong, E. Mok, L. Xia, C. Choi, M. Pyeon, and J. Heo. *Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR*. Automation in Construction, 20(1):3–13, 2011.

[56] C. Yang and H.-R. Shao. *WiFi-based indoor positioning*. Communications Magazine, IEEE, 53(3):150–157, 2015.

[57] A. M. Abusara. *Indoor Positioning Techniques and Approaches for WI-FI Based Systems*. PhD thesis, American University of Sharjah, 2015.

[58] P. Bahl and V. N. Padmanabhan. *RADAR: An in-building RF-based user location and tracking system*. In INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, volume 2, pages 775–784. Ieee, 2000.

[59] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg. *Compass: A probabilistic indoor positioning system based on 802.11 and digital compasses*. In Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization, pages 34–40. ACM, 2006.

[60] *Ekahau*. http://www.ekahau.com/, 2016. [Online; accessed 05-2016].

[61] U. Schatzberg, L. Banin, and Y. Amizur. *Enhanced WiFi ToF indoor positioning system with MEMS-based INS and pedometric information*. In Position, Location and Navigation Symposium-PLANS 2014, 2014 IEEE/ION, pages 185–192. IEEE, 2014.

[62] E. Mok and G. Retscher. *Location determination using WiFi fingerprinting versus WiFi trilateration*. Journal of Location Based Services, 1(2):145–159, 2007.

[63] P. Bolliger. *Redpin-adaptive, zero-configuration indoor localization through user collaboration*. In Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments, pages 55–60. ACM, 2008.

[64] A. Genco. *Three step bluetooth positioning*. In Location-and Context-Awareness, pages 52–62. Springer, 2005.

[65] L. Pei, R. Chen, J. Liu, T. Tenhunen, H. Kuusniemi, and Y. Chen. *Inquiry-based bluetooth indoor positioning via rssi probability distributions*. In Advances in Satellite and Space Communications (SPACOMM), 2010 Second International Conference on, pages 151–156. IEEE, 2010.

[66] Y. Wang, X. Yang, Y. Zhao, Y. Liu, and L. Cuthbert. *Bluetooth positioning using RSSI and triangulation methods*. In Consumer Communications and Networking Conference (CCNC), 2013 IEEE, pages 837–842. IEEE, 2013.

[67] *Indoo.rs*. http://indoo.rs/bluetooth-low-energy-solution-released/, 2016. [Online; accessed 05-2016].

[68] *Blooloc*. https://www.blooloc.com/, 2016. [Online; accessed 05-2016].

[69] *InfSoft*. http://www.infsoft.com/blog/2015/indoor-navigation-and-indoor-positioning-using-bluetooth, 2016. [Online; accessed 05-2016].

[70] M. Sugano, T. Kawazoe, Y. Ohta, and M. Murata. *Indoor localization system using RSSI measurement of wireless sensor network based on ZigBee standard*. Target, 538:050, 2006.

[71] G. Mao, B. Fidan, and B. D. Anderson. *Wireless sensor network localization techniques*. Computer networks, 51(10):2529–2553, 2007.

[72] C. Savarese, J. M. Rabaey, and J. Beutel. *Location in distributed ad-hoc wireless sensor networks*. In Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on, volume 4, pages 2037–2040. IEEE, 2001.

[73] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu. *Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks*. Signal Processing Magazine, IEEE, 22(4):70–84, 2005.

[74] L. Taponecco, A. D'Amico, and U. Mengali. *Joint TOA and AOA estimation for UWB localization applications*. Wireless Communications, IEEE Transactions on, 10(7):2207–2217, 2011.

[75] C. Zhang, M. J. Kuhn, B. C. Merkl, A. E. Fathy, and M. R. Mahfouz. *Real-time noncoherent UWB positioning radar with millimeter range accuracy: Theory and experiment*. IEEE Transactions on Microwave Theory and Techniques, 58(1):9–20, 2010.

[76] K. He, Y. Zhang, Y. Zhu, W. Xia, Z. Jia, and L. Shen. *A hybrid indoor positioning system based on UWB and inertial navigation*. In Wireless Communications & Signal Processing (WCSP), 2015 International Conference on, pages 1–5. IEEE, 2015.

[77] E. Garcia, P. Poudereux, A. Hernandez, J. Urena, and D. Gualda. *A robust UWB indoor positioning system for highly complex environments*. In Industrial Technology (ICIT), 2015 IEEE International Conference on, pages 3386–3391. IEEE, 2015.

[78] G. Shi and Y. Ming. *Survey of Indoor Positioning Systems Based on Ultra-wideband (UWB) Technology*. In Wireless Communications, Networking and Applications, pages 1269–1278. Springer, 2016.

[79] *Pozyx*. https://www.pozyx.io/, 2016. [Online; accessed 05-2016].

[80] *EVARILOS*. http://www.evarilos.eu, 2016. [Online; accessed 05-2016].

[81] T. Van Haute, E. De Poorter, J. Rossey, I. Moerman, V. Handziski, A. Behboodi, F. Lemic, A. Wolisz, N. Wiström, T. Voigt, et al. *The evarilos benchmarking handbook: Evaluation of rf-based indoor localization solutions*. In 2e International Workshop on Measurement-based Experimental Research, Methodology and Tools (MERMAT 2013), pages 1–6, 2013.

[82] T. Van Haute, E. De Poorter, F. Lemic, V. Handziski, N. Wirstrom, T. Voigt, A. Wolisz, and I. Moerman. *Platform for benchmarking of RF-based indoor localization solutions*. Communications Magazine, IEEE, 53(9):126–133, 2015.

[83] F. Lemic, V. Handziski, A. Wolisz, T. Constambeys, C. Laoudias, D. Zeinalipour-Yazti, C. Panayiotou, S. Adler, S. Schmitt, H. Will, et al. *EVARILOS Open Challenge: Track 3*. Telecommunication Networks Group, Tech. Rep. TKN-14-005, 2014.

Part I

Optimizing Indoor Localization Solutions

2

Optimizing Time of Arrival Localization Solutions for Challenging Industrial Environments

In this chapter, we present an optimization of a Time of Arrival (ToA) solution in a challenging industrial environment. In combination with Appendix A, this research was a head start to get familiar with the state of the art. This chapter proofs indoor localization solutions can be "tweaked" for specific environments. As such, the obtained results are not representable for all types of environments making current comparisons between multiple solutions infeasible. It reveals a more standardized methodology is indispensable if an objective comparison is desirable. Nevertheless, the presented results proof a decent calibration and filtering of the measurement data can improve the accuracy results definitely. This optimization process is certainly an added value in environments where multipath fading may occur.

⋆ ⋆ ⋆

T. Van Haute, B. Verbeke, E. De Poorter, I. Moerman.
Published in IEEE Transactions on Industrial Informatics,
April 2016. [IF: 8.785 , Q1]

Abstract Since GPS-technologies cannot be used indoors, a significant amount of research focuses on developing radio-frequency (RF) based alternatives for indoor localization. Unfortunately, most of the suggested solutions for indoor localization consist of theoretical work or have been evaluated in non-industrial environments, typically office spaces. To evaluate the influence of industrial environments on localization accuracy, in this chapter a Time of Arrival (ToA) approach was used to determine the stationary locations of a robot inside the w-iLab.t II testbed, an open industrial-like environment containing several metal obstacles. The ToA method utilizes the measured propagation time of a radio wave between a sender and receiver to estimate their corresponding distance. The chapter evaluates several industrial related deployment aspects that influence location accuracy and describes how their negative impact can be reduced resulting in an almost 50% accuracy improvement in industrial environments.

2.1 Introduction

In recent years, the research focusing on indoor localization solutions has grown exponentially [1–3]. However, as stated in [4, 5], due to the complexity of evaluating indoor solutions at scale, many localization solutions remain either theoretical or are evaluated in easily-accessible environments. Typically, these environments are limited to office, university or home environments [6, 7]. Since radio signals are strongly influenced by the environment, it has been experimentally shown [7] that the accuracy of radio frequency based indoor localization depends strongly on the evaluation environment. Especially the presence of metal obstacles has been shown to strongly influence the propagation behaviours due to effects such as multi-path fading causing reflections. As such, there is a strong need for validating and optimizing current localization solutions in challenging environments. To this end, this chapter evaluates a Time of Arrival (ToA) localization solution in an industrial-like wireless testbed, which consists of a large open area with metal walls and ceilings, as well as several metal objects such as large horizontal and vertical metal pipes.

Time of Arrival localization solutions estimate distances between devices based on the propagation time of a Radio Frequency (RF) wave between sender and receiver. Since the speed of propagation waves is known, a corresponding distance can be determined based on the measured time. When *one way* ToA is used, a very accurate clock synchronization is required between sender and receiver. Providing such accurate synchronization is difficult, especially on low-end devices [8, 9]. However, this can be circumvented by using *two way* ToA-measurements. In this case, the single Round-trip Time (single RTT) is measured and clock synchronization is no longer necessary. Even more accurate results can be obtained when using symmetrical-double sided two-way ranging (SDS-TWR) [10] which further improves the accuracy by sending one additional message per ranging measurement. However, the SDS-TWR approach assumes that the fixed nodes also perform and collect ranging measurements, and requires communication links between the anchor nodes to distribute this information, making the SDS-TWR approach more complex and expensive to install.

The results described in this chapter are obtained using single RTT. The distance between two nodes using *two way* ToA can be determined using the following equation:

$$\begin{aligned} d &= \frac{c}{2} t_{flight} \\ &= \frac{c}{2} (t_{RTT} - t_{delay}) \end{aligned} \tag{2.1}$$

With c the speed of light, t_{flight} the time it takes for the signal to propagate, t_{RTT} the measured RTT between sender and receiver and t_{delay} the time used for processing the packets. These processing delays include the time required (i) to send the packet (ii) to process the received packet and (iii) to generate an Acknowledgement (ACK).

t_{flight} is the time measured between transmission of the entire packet and reception of the entire ACK. It is important to measure t_{delay} as precisely as possible so that t_{flight} depends only on the distance. Both t_{RTT} and t_{delay} are measured in number of clock ticks and can be converted to time using the formula $\frac{n_{cycles}}{f_{timer}}$. Hereby, n_{cycles} is the number of measured clock ticks and f_{timer} is the frequency of the used timer. The resolution of one clock tick permits a spatial precision that is equal to:

$$\Delta_d = \frac{c}{2 f_{timer}} \tag{2.2}$$

Based on this equation it can be concluded that the higher the clock frequency, the higher the accuracy of the estimated distance. In this chapter, a constrained device with a clock speed of 12 MHz is used which is a typical clock speed for a cheap embedded processor. The corresponding spatial resolution is 12.5 m or, in other words, one clock tick equals 12.5 m. However, by executing multiple ranging measurements, it is possible to achieve a sub-clock precision [11]. Once distances from the mobile node to the anchors are calculated, i.e. the “ranging” process, these distances can be used to estimate the position of the mobile node.

The main contributions of this chapter are as follows. (i) A ToA localization solution is evaluated in an “industrial-like” environment, i.e. an open space with metal obstacles, causing a significant amount of reflections and multipath fading. (ii) The impact of several deployment aspects (height, antenna orientation, the presence of external interference, etc.) on the ranging accuracy are analysed. (iii) Different filtering and positioning estimators are evaluated to derive a localization solution that is suitable for harsh environments.

The remainder of this chapter is structured as follows. Section 2.2 describes related work. In Section 2.3, a detailed overview of the evaluation environment is given. Afterwards, Section 2.4 discusses deployment aspects for indoor localization solutions. Next, Section 2.5 discusses solutions for removing outlier values caused by multi-path fading in industrial environments. Two different location algorithms are discussed in Section 2.6. Finally, Section 2.7 concludes the chapter.

2.2 Related work

Currently, most indoor localization solutions utilize RF based devices due to their low cost [6]. Typically, many RF based solutions utilize IEEE 802.15.4 based wireless sensor nodes due to their low cost and energy requirements, and because their low layer behaviour can easily be modified. These sensor nodes are equipped with clocks with a low frequency, which according to equation (2.2) results in inaccurate distance estimates.

For the round trip time calculations, typically the Acknowledgements (ACKs) mechanism is exploited, originally designed for increasing the reliability of packet transmissions. Due to the unreliable nature of wireless communication, there is no guarantee that transmitted packets will arrive at the receiver. Therefore, the ACK-principle was used to reply each received packet with an acknowledgement. Because IEEE 802.15.4 has very strict timing specifications, the ACK-mechanism is often built into the hardware of the radio chip and as such the t_{delay} can be measured very precisely. In [12] a comparison between software and hardware ACKs was done for varying frame sizes. The conclusion was that hardware ACKs are not sensitive to different frame sizes and they are very stable in terms of timing. By contrast, the software ACKs exhibit some variation in time. Therefore, typically hardware ACKs are chosen in order to determine the number of clock ticks as accurate as possible.

In [9, 11], the importance of obtaining accurate timing information is shown, which in turn depends on the accuracy of the oscillator. In [9], a comparison of the crystal oscillator and the DCO (Digital Controlled Oscillator) is made. The DCO is strongly influenced by the temperature. For a certain distance, a difference of 30 clock ticks is measured when the temperature raises 30 degrees. If the oscillator has a clock frequency of 12 MHz, this corresponds with a distance deviation of up to 375 meter. As such, the experiments in this chapter utilize a crystal oscillator.

Several studies exist that evaluate localization solutions in harsh environments [1, 4, 7, 13–17]. However, most of these utilize fingerprinting solutions and realize low accuracies. A smaller selection of papers [7, 14–17] evaluate a ToA solution in an industrial-like environment. Besides demonstrating that ToA solutions suffer from degraded performance in these challenging environments, only [15, 16] propose a solution to improve the accuracy. Of those, [15] achieves similar accuracies as the solution proposed in this chapter, but the results are only validated using simulations and in contrast to our solution pre-defined, fixed anchor positions are required thereby limiting deployment flexibility. A real-time self-calibration ToA solution is proposed in [16]. However, in contrast to our solution this requires continuous measurement exchanges among the anchors in the wireless network, thereby imposing the need for a reliable, low-delay communication infrastructure between the anchor points, again limiting deployment flexibility.

2.3 Industrial environment

2.3.1 Wireless testbed facility

The w-iLab.t II testbed [1] is located in Zwijnaarde, a city near Ghent, above a cleanroom. The testbed consists of a large 66 m x 21 m area with metal walls, metal ceiling and a metal floor. Distributed over the area are several horizontal and vertical metal pipes, illustrated in Figure 2.1 and as such the area represents a challenging industrial like environment with significant multi-path fading effects causing reflections. The environment is shielded from outside interference, but minimal levels of external interference might leak through. This external interference originates from outside access points, can only very sporadically be observed at the borders of the test area and has low energy values. This results in an environment with no uncontrollable interference that can influence the results and where thus very reproducible wireless experiments can be performed.

The presence of multipath fading and thus reflections in this environment is significantly higher than in a typical office environment. To illus-

[1] http://ilabt.iminds.be

trate this fact, Figure 2.2 shows the outcomes of a channel sounder experiment [18] in a typical brick-wall office testbed and the w-iLab.t II testbed. During this experiment, a signal is transmitted and the reflections are captured. In contrast to the office environment (Figure 2.2a), the w-iLab.t II testbed (Figure 2.2b) is characterized by a large number of unpredictable, high strength reflections, causing challenging conditions for indoor localization [6, 19].

Figure 2.1: The w-iLab.t II wireless testbed: an open environment with metal obstacles. This test environment has similar specifications as a challenging industrial environment.

2.3.2 Hardware devices

Over an experimentation area, 60 fixed nodes with a known location are distributed. In addition, mobile nodes are available that can move around for localization experiments. These nodes are based on a vacuum cleaning robot and are extended with a radio for remote control and accurate positioning algorithms. Due to the fact that the movement of these robots is controlled, mobility is reproducible. The fixed nodes are marked with blue spots whilst the mobile nodes have orange spots on the map in Figure 2.3.

The mobile robots can transmit packets, which are acknowledged by

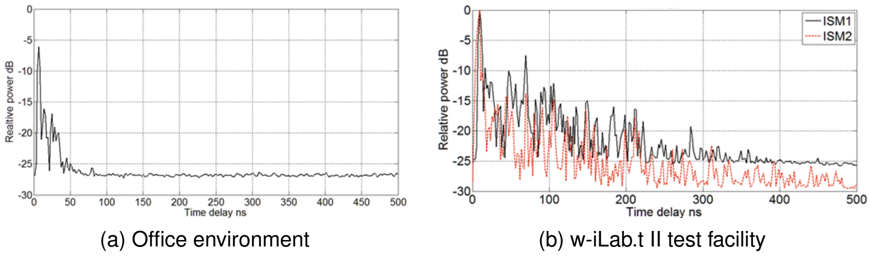


Figure 2.2: Channel sounder experiments in an office environment and in the w-iLab.t II environment. In the office environment reflections are minimal since the walls absorb the signals. In the w-iLab.t II test facility the presence of metal obstacles causes multipath effects such as reflections. Two different frequency bands are evaluated in the Industrial, Scientific and Medical (ISM) radio bands. The ISM1 band is located in 2.2-2.75 GHz whilst the ISM2 band is located in 4.4-5.9 GHz [18].

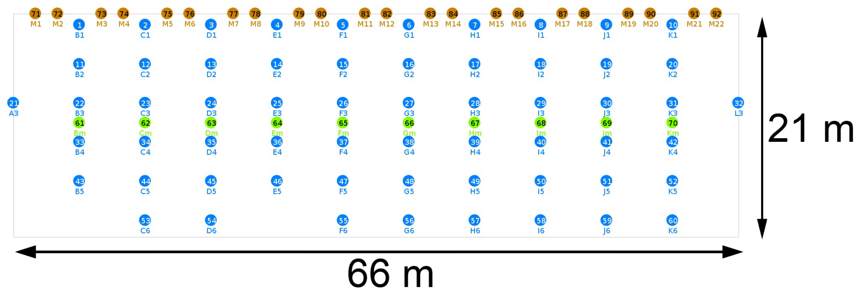


Figure 2.3: The w-iLab.t II wireless testbed node map inlcuding 60 fixed nodes (blue) and 20 mobile robots (orange) [7].

the fixed nodes. De default transmission power is 8 dBm. The RTT is used for calculating the distance to each anchor point using equation (2.1). In all our experiments, Rmoni sensor devices were used as the receiver on the fixed locations, whilst the STM32W-RFCKIT [2] functions as the sender on the mobile robot. This microcontroller contains four oscillators: (i) a high frequency RC oscillator (HSI), (ii) a 24 MHz crystal oscillator (HSE), (iii) a 10 kHz LSI RC oscillator (LSI10K) and (iv) a 32.768 kHz crystal oscillator (LSE). The default system clock (SCLK) uses the HSI oscillator and is mainly used on the STM32 board. However, the radio chip uses the HSE clock because strict timing specifications are required for measuring the ACKs.

2.4 Deployment aspects

These sections evaluate the impact of several deployment aspects on the expected accuracy.

2.4.1 Influence of the height on the packet loss rate

Several industrial or warehouse use cases envision the use of small robots for moving equipment, for example for material handling by automated guided vehicles. Since these robots are often very compact, the use of radio solutions can be impacted by the distance to the ground. The first experiment investigates the relation between the height of the sensor nodes and the packet loss rate. The packet loss rate PL can be defined as:

$$PL = \frac{r}{t} * 100 \tag{2.3}$$

Whereby r is the number of requests that did not receive an ACK and t the total number of requests. The setup of this experiment was the following: the receiver was placed 10 cm above the ground during the entire experiment. The sender was placed on two positions: on the floor or 50 cm above the floor. For each position, multiple distances were used in order to measure the drop rate. At 5 m, the packet loss rate was as high as 99 % when the sender was located on the floor. For a height of 50 cm, the packet loss rate is much better and can even decrease with increasing distances due to reflections on the ground [11]. All intermediate results can be found in Table 2.1. These results indicate clear limits to the minimum

[2] Datasheet: http://www.st.com/web/en/resource/technical/document/data_brief/CD00296406.pdf

height of a robot intending to be used for RF based localization solutions in, for example, automated warehouses [20].

Table 2.1: Packet loss rates for different sensor heights (0 cm and 50 cm) and different distances. Deploying the sensor nodes nearby the floor significantly increases the packet loss rates.

Distance [cm]	0 cm [%]	50 cm [%]
10	11.80	14.27
100	18.42	19.12
200	17.41	19.30
300	34.82	19.31
400	45.64	14.96
500	**99.17**	**17.63**

2.4.2 Influence of the antenna orientation

The transmission power of an antenna is not equally distributed in each direction [21]. The radiation pattern of a typical omni-directional antenna is shown in Figure 2.4. Since the STM-32W is using a printed meandered monopole antenna, it is most-likely polarized linearly. The linear polarization is in the same plane as the PCB as shown in Figure 2.5. The two subsections below describe respectively the influence of the mutual direction and the absolute angle of the antenna.

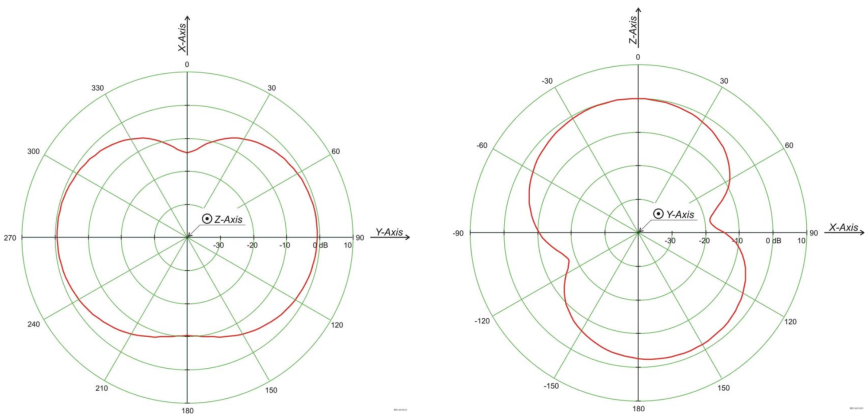


Figure 2.4: Antenna radiation pattern of the STM32W-RFCKIT in the XY and XZ pane. [21].

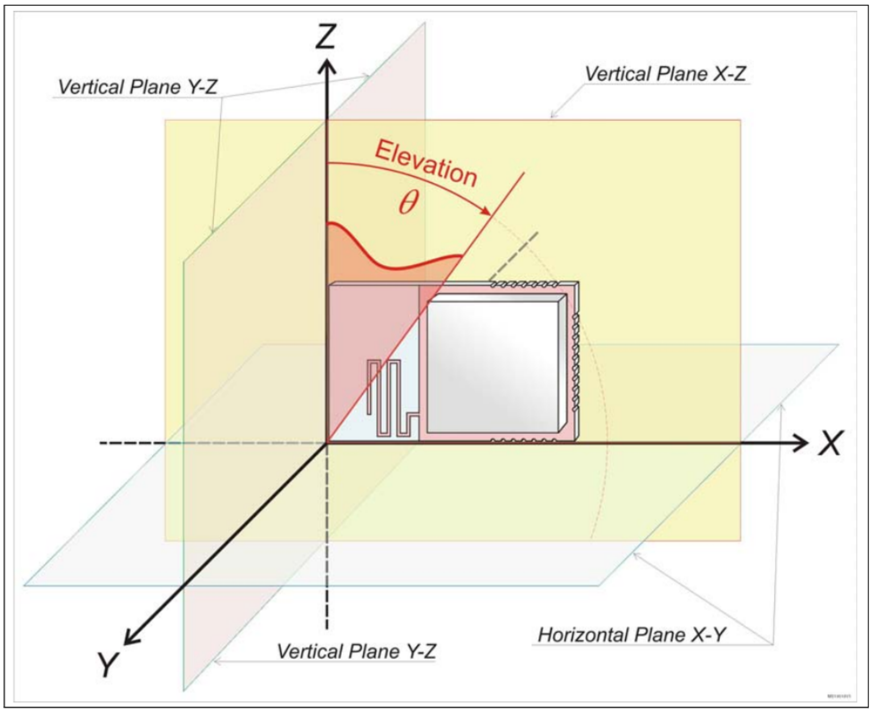


Figure 2.5: Visualization of the PCB with the printed meandered monopole antenna. [21].

2.4.2.1 Influence of the mutual direction

In order to investigate the influence of the antenna orientation, both sender and receiver were placed on a height of 50 cm. Next, they were placed on multiple distances and in three different directions. The used directions were:

- **S:H R:H** - Both **s**ender and **r**eceiver were placed **h**orizontally.
- **S:H R:V** - The **s**ender was placed **h**orizontally, the **r**eceiver **v**ertically.
- **S:V R:V** - Both nodes were placed **v**ertically.

For each distance and orientation, the data was collected for five minutes. The results are shown in Table 2.2, each direction is evaluated with three different distances: 3 m, 9 m and 15 m. To verify the stability of the conclusions, several experiments were executed twice at different moments.

Table 2.2: The results of the different mutual orientations between sender and receiver are listed. The figures show that a different antenna orientation significantly influence the ranging accuracy (up to 4.52 clock ticks or 56.5 m).

Distance [m]	Measured number of clock ticks		
	S:H R:H	S:H R:V	S:V R:V
3	3812.97	3811.17	3813.58
	3812.64	3810.54	
9	3818.42	3814.86	3813.90
	3817.94		3814.52
15	3814.57	3816.47	3813.84
			3813.34

The measurements that were repeated with the same direction and distance, show small differences. A maximum difference of 0.63 clock ticks is registered. This corresponds to 7.88 m using Equation (2.2). Besides that, the measurements on the same distance with a different orientation show a much bigger difference: maximum 4.52 clock ticks (or 56.5 m). Since the antennas do not send their signals with an uniform strength in all directions (Figure 2.4) due to multipath reflections the strongest signal can have a different path when the orientation of the transmitter changes. This confirms the fact that practical deployments need to take care to orient all antennas in the same direction if possible.

2.4.2.2 Influence of the antenna angle

For this experiment, the fixed nodes are mounted vertically on the rods shown in Figure 2.6 and the transmitter on the mobile robot is positioned horizontally. Because the robot can rotate around its own axis, the difference in clock ticks when the angle changes between the mobile and fixed node can be measured. The robot was placed in the centre of four different fixed nodes (Node 44, 45, 53 and 54). Next, the robot rotated 45 degrees 8 times. After each rotation, the robot stayed in this position for collecting data during a period of 30 seconds. An overview of the results of this experiment can be found in Table 2.3, Table 2.4 shows the more detailed results for one fixed node. Turning the device 45 degrees can result in an distance difference of 57.2 m, with worst case difference (node 45) up to 110.88 m.

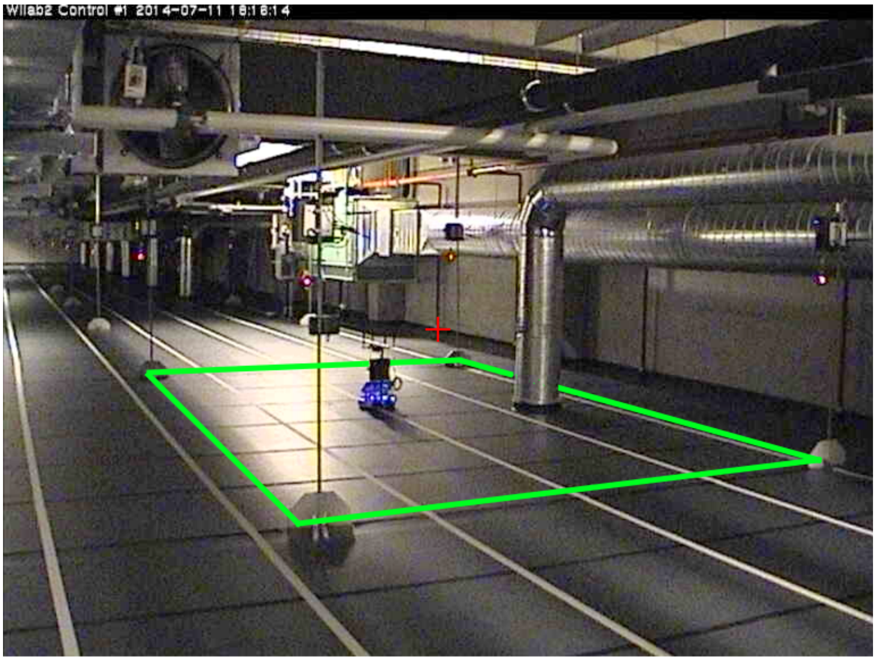


Figure 2.6: Setup for the investigation of the influence of the angle.

2.5 Optimization of the ranging process

In many RF based localization systems, location estimates first require a “ranging process” during which the distance between the mobile device

Table 2.3: Impact of the horizontal antenna orientation of the sensor node on ToA measurements. Rotating the mobile device can influence the range estimations by up to 110.88 m.

Node	Min clock ticks	Max clock ticks	Max difference clock ticks	Max difference distance [m]	Stdev clock ticks
44	3811.43	3816.22	4.79	59.88	1.64
45	3807.46	3816.33	8.87	110.88	2.69
53	3810.62	3813.64	3.02	37.75	1.11
54	3812.12	3814.53	2.42	30.25	0.83

Table 2.4: The measured clock ticks between the mobile node and fixed node 44 for each angle of the robot. Turning the robot 40° can result in a difference of almost 4 clock ticks.

Node	Orientation angle [°]	Measured clock ticks
44	145	3816.22
	93	3811.64
	47	3813.81
	4	3813.87
	-43	3811.43
	-85	3815.36
	-134	3813.12
	-175	3813.33

(with unknown location) and each fixed node (with known location) is estimated. This section experimentally analyses how the uncertainties of the ranging process, caused by challenging industrial environments, can be reduced.

2.5.1 Influence of outliers

The previous subsections demonstrate the sensitivity of the measurements due to the environment and the mutual orientation of the sensor nodes. In this section, filter techniques are proposed to improve the robustness against the issues of the hardware orientation. For this experiment, 22 measurement points were selected over the w-iLab.t II testbed area. Ranging measurements were collected for one minute at each measurement point, 35 fixed nodes were activated during these measurement campaign. In total, 245.000 ToA measurements were collected, stored and processed.

To demonstrate the importance of removing outliers, this section will first calculate the location accuracy with the unfiltered data. The processing delay (t_{delay} of Equation (2.1)) was initially determined by measuring the time when two nodes were placed at a negligible distance ($<$ 15 cm) from each other. Since t_{delay} is known, the distance can be calculated using Equation (2.4).

$$d = \frac{c(t_{RTT} - t_{delay})}{2f_{timer}} \tag{2.4}$$

At the different measurement points, the estimated distances were compared with the actual distances to the fixed nodes (i.e. the "ranging error"). The results of these ranging errors are shown in a CDF (Figure 2.7). This shows that in 50 % of the cases, the ranging error is smaller or equal to 31.99 m. The corresponding Root Mean Square Error (RMSE), calculated using Equation (2.5), corresponds to 277.78 m. These inaccuracies are too large to be used as input for location estimates.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{d}_i - d_i)^2}{N}} \tag{2.5}$$

Due to (self-)interference, several anchor points have lossy links with the device that needs to be located. Since these anchor points provide only a limited number of responses, we do not consider these anchor points reliable. By removing all ranging measurements from anchor points whereby the amount of received replies is lower than 50, the average ranging error is halved. Next, we analyze the distribution of the remaining anchor points.

In Figure 2.8, a histogram of one node is shown. This histogram shows that the majority of estimated clock ticks falls within a range of only three values. The other values (in the region of 3800 and 3870) will influence the ranging error despite their small number. To remove these outliers, a median and standard deviation filter technique is investigated, applied and compared. These filter techniques are also applied in [9] and [11]:

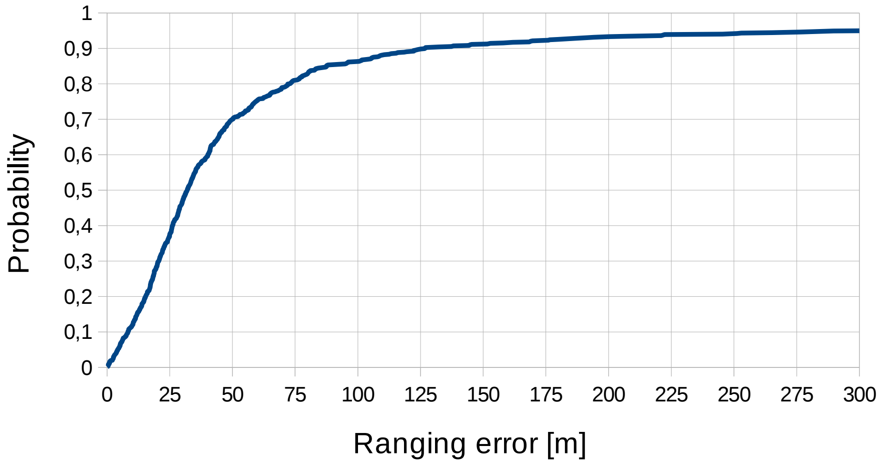


Figure 2.7: CDF of the ranging errors using unfiltered measurements.

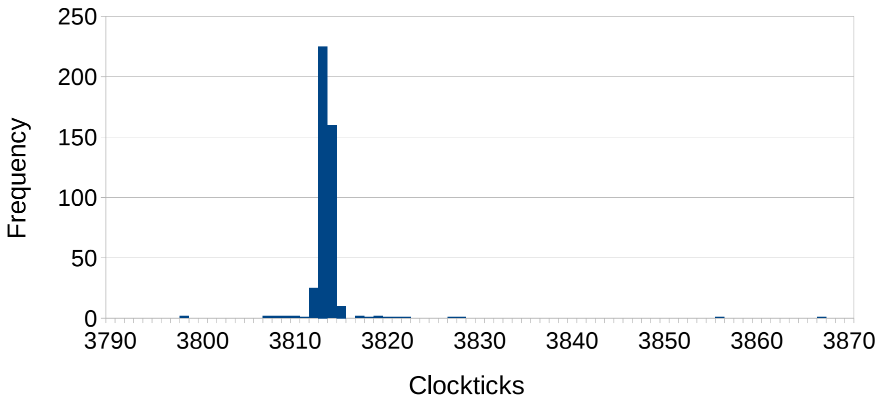


Figure 2.8: The ranging results between the mobile node and one fixed node (node 44) is represented in a histogram. This shows a typical spreading of the ToA measurements.

- Median filtering: this filter removes the 10% highest and lowest measurements. Then the median is calculated and all measurements above or below a certain threshold, are removed.

- Standard deviation filtering: this technique calculates the standard deviation of the unfiltered measurement results. Subsequently, all measurements that are higher or lower than one time the standard deviation are removed.

No motivation was given in [9, 11] why these specific limits (10% and 1 * stddev) were applied. To evaluate the optimal limits for our test environment, the optimal parameter values were determined experimentally. The schematic representation of the process can be found in Figure 2.9.

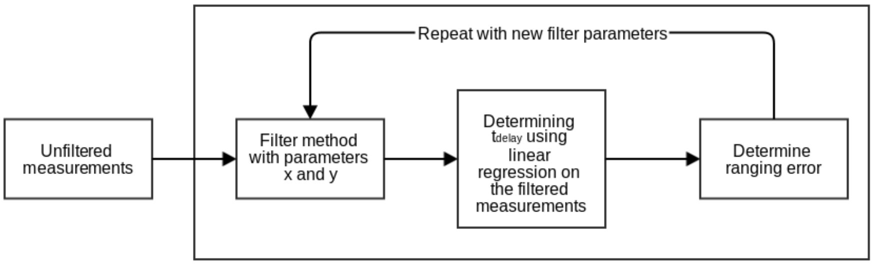


Figure 2.9: The used filtering methodology for determining the optimal filter limits.

2.5.1.1 Median filtering

Applying the median filter as described in the previous section, gave positive results. The average error was reduced to 15.64 m and the RMS error after filtering is reduced to 21.70 m. The percentage limits for removing the highest and lowest measurements did not affect the new median since the median is known to be a very robust statistic and the number of outliers is very small. For those reasons, the 10% was used similar to [9, 11].

2.5.1.2 Standard deviation filtering

The results of the standard deviation filtering technique for different filter limit values can be found in Figure 2.10. The ranging errors are lowest when a value between 0.25% and 0.5% is used.

2.5.1.3 Summary

A summary of both filter techniques can be found in Table 2.5 whereby the average, RMS and median error distance are listed. Both filtering methods result in similar improvements of the ranging errors, increasing the average accuracy of range estimates from 106 meter to around 17 meter.

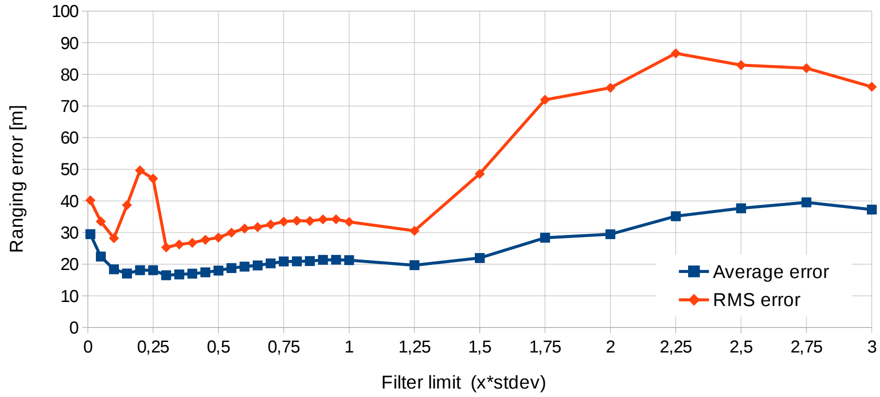


Figure 2.10: The average and RMS ranging error in function of multiple filter limits

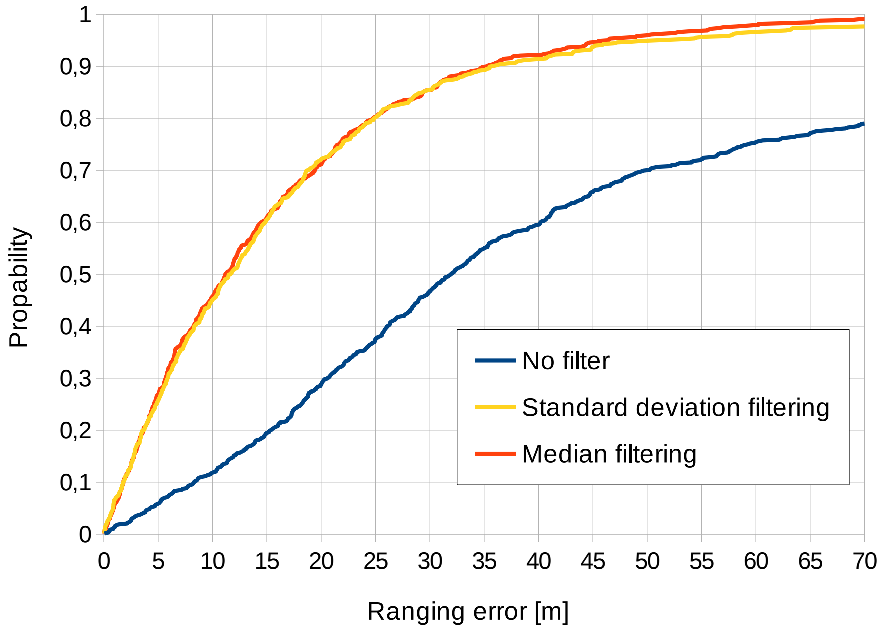


Figure 2.11: CDF showing the positive influence of both filter methods on the ranging error.

Table 2.5: Reduction of ranging errors by applying two different filter techniques: median and standard deviation filtering. Both solutions achieve similar improvements.

Filter	# ACK's	% of total	Ranging error Average [m]	RMS [m]	Median [m]
Unfiltered	158 870	100	106.33	323.38	38.82
Median	125 006	78.68	17.57	23.56	13.56
Std. dev.	136 721	86.06	17.65	23.75	13.44

2.5.2 Influence of the increasing distance on the ranging errors

In this section we want to formulate an answer to the following question: which measurements are the most reliable? Those from the nearby anchor nodes or those which are further away from the mobile node? The results of this experiment can be found in Table 2.6. The distance estimations are divided in intervals of 5 m. For each interval, the average and RMSE distance were calculated. These are the results after the median filter technique was applied.

Table 2.6: Distance between nodes and the corresponding ranging error.

Actual distance [m]	# distance estimations	Ranging error Average [m]	RMS [m]
$[1-5[$	25	6.46	8.56
$[5-10[$	118	8.83	11.66
$[10-15[$	135	9.72	11.92
$[15-20[$	125	9.85	12.23
$[20-25[$	83	10.54	12.19
$[25-30[$	54	10.82	13.04
$[30-35[$	31	11.44	13.74
$[35-40[$	21	8.36	10.11
$[40-45[$	9	5.16	6.85
$[45-50[$	1	4.80	4.80
$[50-55[$	3	6.23	7.20

The results show that the ranging error initially increases with distance, but starting from 40 to 55 meters starts decreasing again. A possible explanation for this behaviour is the following. When nodes are close to each other, they are often within Line Of Sight (LOS) distance, resulting in more predictable behaviour and hence lower ranging errors. When the distance increases, links become more unpredictable due to multipath effects resulting in higher errors. When increasing the distance even more, the received signal strengths decrease and as a result less packets are received (as shown in Table 2.6) but the packets that are received are likely within LOS and thus result in lower ranging errors. As a result, at higher distances less links can be used for communication, but the ones that are received are more accurate.

As a solution to use mainly the more predictable data packets from nearby nodes the output power can be lowered. Table 2.7 shows the ranging accuracy for an experiments in which the output power is reduced from 8 dBm to -19 dBm, further reducing the average ranging error by almost 50% from 17 m to 9 m. As such, the overall filtering process (median filtering + reducing the output power) reduces the average ranging errors from 106 m to around 9 m.

Table 2.7: Influence of decreasing the output power on ranging errors.

Filter	# ACK's	% of total	Ranging error Average [m]	RMS [m]	Median [m]
Unfiltered	28 369	100	42.13	110.52	29.16
Median	21 905	77.21	9.26	11.61	7.62
Std. dev.	23 240	81.92	9.52	11.98	8.05

2.5.3 Influence of external interference

Finally, in many realistic industrial environment external interference from e.g. Wi-Fi or other IEEE 802.15.4 networks is unavoidable. The impact of interference on localization solutions has only recently become a topic of study. In [22] interference was artificially generated in office environment to demonstrate that some types of localizations solutions exhibit degraded performance in the presence of interference whereas other solutions showed minimal impact. However, [22] did not investigate ToA solutions and to the best of our knowledge, no information is available about the impact of interference on localization experiments specifically in industrial environments. To investigate this aspect in more detail, experiments with varying levels of interference were performed. Interference was generated by sending traffic between Wi-Fi nodes that were located at a distance of around 2 m from the mobile node and used a transmission power of

20 dBm. Up to two interference generating Wi-Fi pairs were used simultaneously. In order to verify the correctness of our interference scripts, a spectrum sensing device registered and visualized the 2.4 GHz spectrum. The Wi-Fi interference sources used channel 1 since our ToA solution uses channel 11. As such, both channels overlap and the interference sources could affect the ranging measurements. The presence of the interference can be confirmed using Figure 2.13. The results of the experiments are shown in Table 2.8. The variations observed during experiments with interference are similar to the expected variations without interference. Although variations in terms of minimum, maximum and average clock tick values can be observed, the filtering approach removed all these effects and is sufficiently robust against even high levels (300 Mbit/s) of external interference.

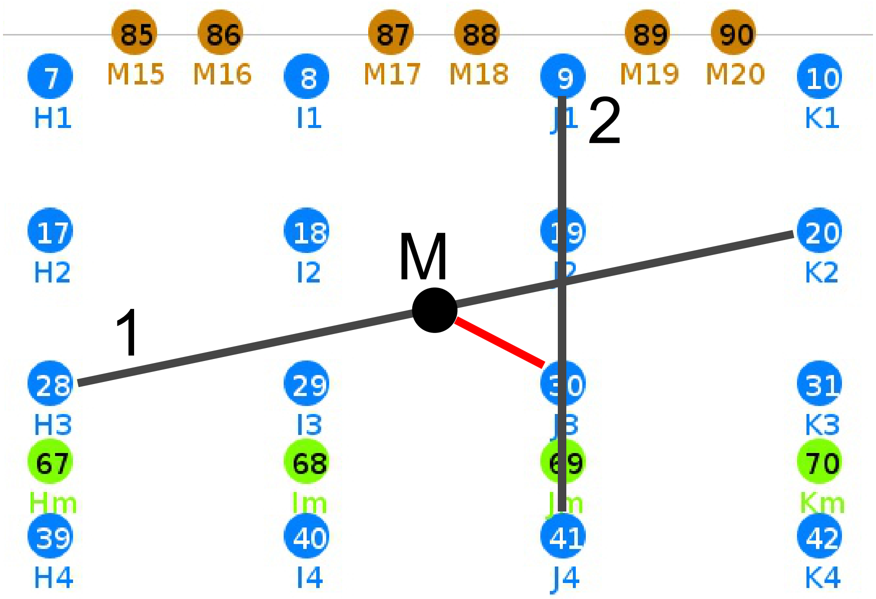
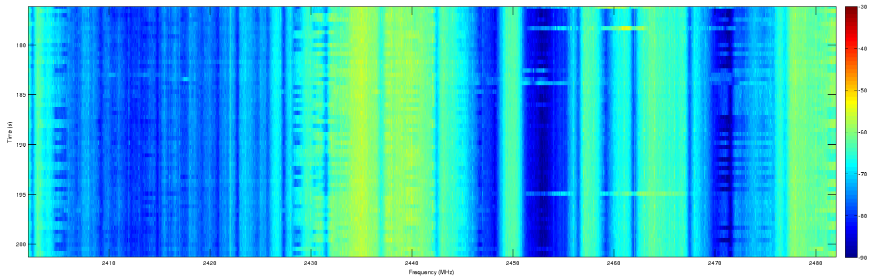
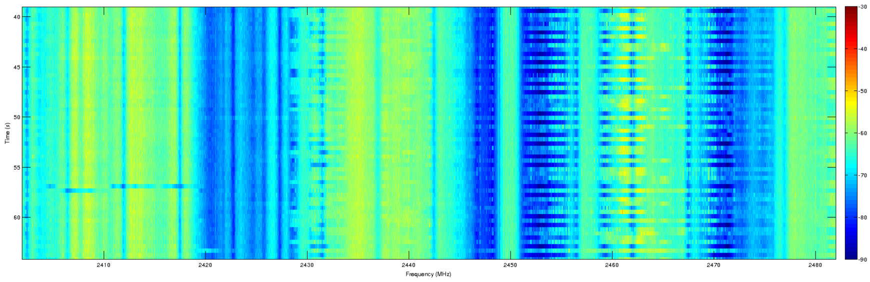


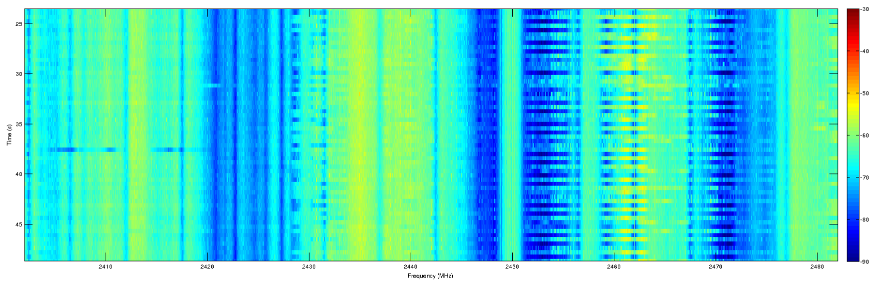
Figure 2.12: Test setup evaluating the influence of interference: M points the location of the mobile node, which was collecting ranging measurements of anchor node 30 (red connection). Additionally, two anchor pairs (grey connections) were used to create Wi-Fi interference.



(a) No interference



(b) Pair 1



(c) Pair 2

Figure 2.13: Plots of the spectrum during the interference experiments. Figure 2.13a confirms no energy was detected when the interference scripts were not executed. Figure 2.13b and 2.13c show a strong activity on channel 1. The reason why 2.13b visualizes a stronger signal is due to the fact that our spectrum sensing device is located at node 17 (Figure 2.12) which is closer to Pair 1 than Pair 2.

Table 2.8: Impact evaluation of interference on the ranging process. The scenarios are sorted according to increasing interference levels.

		Clock ticks			
ID	Description	Min	Max	Mean	Median
1	No interference	5139	5846	5750	5780
9	No interference	5203	5846	5765	5780
6	No interference	5200	5846	5764	5780
2	Pair 1 (30 Mbit/s)	5191	5845	5762	5780
7	Pair 2 (30 Mbit/s)	5200	5845	5765	5780
3	Pair 1 & 2 (30 Mbit/s)	5199	5846	5750	5780
5	Pair 1 (300 Mbit/s)	5203	5845	5764	5780
8	Pair 2 (300 Mbit/s)	5181	5846	5750	5780
4	Pair 1 & 2 (300 Mbit/s)	5195	5845	5750	5780

2.6 Localization algorithms

After obtaining the estimated distance to each fixed device, all ranging information is combined to obtain an estimate of the current position of the mobile device. Two different algorithms are evaluated: maximum likelihood and min-max.

2.6.1 Used localization algorithms

2.6.1.1 Maximum likelihood

This method [23, 24] is based on a cost function that is minimized to determine the position of the mobile node. In our case, the cost function can be defined as follows:

$$(\hat{x}, \hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{i=1}^{N} \left(\sqrt{(x - x_i)^2 + (y - y_i)^2} - \hat{d}_i \right)^2 \tag{2.6}$$

Hereby, $(\hat{x}, \hat{y})$ is the estimated position of the mobile node. In the equation above, only the x and y are unknown. These values need to be determined in such a way that the function is minimal. This equation can be classified as a least square optimization problem which can be solved in two ways. The first way is the analytic method like the Newton-Gauss algorithm. The second way uses a grid technique whereby the cost function is calculated on each dot of the grid. The first approach is faster, but includes the possibility that the local optimum is reached instead of the global optimum.

2.6.1.2 Min-max

The algorithm “Min-max” [25, 26] starts with drawing a square around the known positions of the fixed nodes. For each fixed node FN_i the centre of the square will be (x_i, y_i). The distance from the centre to the middle of the side will be defined by the distance estimation $\hat{d}_i$. The square can be described as follows:

$$\left[x_i - \hat{d}_i, y_i - \hat{d}_i\right] \times \left[x_i + \hat{d}_i, y_i + \hat{d}_i\right] \tag{2.7}$$

The next step is determining the intersection of all these squares by finding the maximum value of the coordinate minima (the bottom left corners) and the minimum value of all the coordinate maxima (the upper right corners):

$$\left[\max(x_i - \hat{d}_i), \max(y_i - \hat{d}_i)\right] \times \left[\min(x_i + \hat{d}_i), \min(y_i + \hat{d}_i)\right] \tag{2.8}$$

Finally, the coordinates of the location estimation are defined by the centre of the intersection:

$$\begin{aligned}(\hat{x}, \hat{y}) = &\left[\frac{\max(x_i - \hat{d}_i), \max(y_i - \hat{d}_i)}{2}\right] \times \\ &\left[\frac{\min(x_i + \hat{d}_i), \min(y_i + \hat{d}_i)}{2}\right]\end{aligned} \tag{2.9}$$

2.6.2 Results

The results of the localization algorithms can be found in Table 2.9. The first columns describes the algorithm in combination with the used transmission power. To evaluate the stability of the solution, experiments are repeated several times. In total, four different datasets are recorded at different times and/or dates. The second column represents the time stamp when the measurement data is collected.

Based on Table 2.9, it is clear that both solutions perform best when the transmission power is reduced from 8 dBm to -19 dBm. The improvement is especially noticeable with the maximum likelihood algorithm. In all cases, the min-max algorithm outperforms the maximum likelihood algorithm. To ensure that the experiments are repeatable, new data was collected 10 and 13 days after the first dataset. The deviation with the original experiment is less than half a meter, demonstrating that the solution is quite stable. The results demonstrate that current localization solutions in

Table 2.9: Location accuracy when using the min-max and maximum likelihood localization algorithms.

Algorithm	Timestamp	Ranging error Average [m]	Median [m]	Min [m]	Max] [m]
Max like 8dBm	09/08/2014 13:32:47	33.13	33.81	8.32	54.31
Max like -19dBm	09/08/2014 17:35:20	4.03	3.33	0.39	9.04
Min-max 8dBm	09/08/2014 13:32:47	7.20	6.72	1.62	14.36
Min-max -19dBm	09/08/2014 17:35:20	**3.26**	**2.82**	**0.14**	**6.94**
Min-max -19dBm	19/08/2014 07:05:53	3.60	3.28	1.81	6.86
Min-max -19dBm	21/08/2014 07:06:52	3.65	3.22	1.47	6.90

industrial environments can expect an average accuracy of about 2 m when correctly calibrated. In contrast, the previously mentioned work [7] evaluated localization solutions in multiple environments, but did not optimize these solutions for the different environments. When using non-optimized localization solutions, [7] obtained an average accuracy of only around 6 m for ToA and RSSI based solutions in w-iLab.t II. As such, it is clear that indoor localization solutions in industrial environments based on RF signals are feasible, on the condition that sufficient calibration and filtering is applied, thereby improving the localization accuracy by a factor of almost three when compared to non-optimized solutions.

2.7 Conclusion

Although indoor localization solutions are relevant for a wide range of industrial applications, the number of solutions that is actually evaluated in industrial-like environments is severely lacking. To remedy this, in this chapter an indoor localization solution based on Time of Arrival was evaluated in a challenging large-scale environment, the w-iLab.t II testbed. The evaluation environment consists of an area of 1386 m^2 with metal walls, ceilings and floors, as well as several metal obstacles such as vertical and horizontal pipes.

The chapter demonstrates through channel sounder experiments that the testbed environment is indeed a challenging environment in terms of wireless behaviour such as multipath fading effects that cause reflections. Experiments indicate several deployment aspects that impact the signal strengths and hence, due to multipath effects, result in uncertainties during the ranging process, most notably the antenna height, the mutual antenna orientation and the distance.

Several techniques were evaluated to cope with the presence of these outliers. By applying both basic statistics and reducing the output power, the ranging errors (i.e. the estimated distance to the fixed nodes) improved by 83.48%. Finally, two algorithms, min-max and maximum likelihood were evaluated with the ToA-data. In our case, the min-max solution performed the best in combination with the low transmit power and the median filtering technique. A comparison of these results with previous evaluations in similar conditions [7], shows an accuracy improvement of 47.40% of the average error distance. The error in 50% of the cases lowered to 2.39 m (an improvement of 43.52%). The minimum error distance is increased with 0.15 m (loss of 15.14%) however the maximum error distance is lowered more than three times (from 27.06 m to 6.90 m, an improvement of 74.51%). To conclude, when using correct filtering methods and localization algorithms localization accuracy can be significantly improved in industrial environments up to around 2 m. While these results are not sufficient for all industrial applications, the obtained accuracy is sufficient for a wide range of relevant industrial use cases such as asset or personal localization.

References

[1] K. Derr and M. Manic. *Wireless Sensor Networks-Node Localization for Various Industry Problems*. Industrial Informatics, IEEE Transactions on, 11(3):752–762, 2015.

[2] D. Zhang, S. Zhao, L. T. Yang, M. Chen, Y. Wang, and H. Liu. *NextMe: Localization Using Cellular Traces in Internet of Things*. Industrial Informatics, IEEE Transactions on, 11(2):302–312, 2015.

[3] J. Pak, C. Ahn, Y. Shmaliy, and M. Lim. *Improving reliability of particle filter-based localization in wireless sensor networks via hybrid particle/FIR filtering*. Industrial Informatics, IEEE Transactions on, 11(5):1089–1098, 2015.

[4] M. Hlzl, R. Neumeier, and G. Ostermayer. *Localization in an Industrial Environment: A Case Study on the Difficulties for Positioning in a Harsh Environment*. International Journal of Distributed Sensor Networks, 2015:11, 2015.

[5] T. V. Haute, E. D. Poorter, F. Lemic, V. Handziski, T. Voigt, A. Wolisz, I. Moerman, et al. *Platform for benchmarking of RF-based indoor localization solutions*. Communications Magazine, IEEE, 53(9):126–133, 2015.

[6] H. Liu, H. Darabi, P. Banerjee, and J. Liu. *Survey of Wireless Indoor Positioning Techniques and Systems*. IEEE Trans. Sys. Man Cyber Part C, 37(6):1067–1080, 2007.

[7] T. Van Haute, E. De Poorter, I. Moerman, F. Lemic, V. Handziski, A. Wolisz, N. Wirström, and T. Voigt. *Comparability of RF-based Indoor Localization Solutions in Heterogeneous Environments: An Experimental Study*. page . Inderscience Publishers, 2015.

[8] T. Van Haute, E. De Poorter, F. Lemic, V. Handziski, N. Wirström, and T. Voigt. *Report on experiments without interference*. In EVARILOS Project Deliverable, 2013.

[9] P. Pettinato, N. Wirström, J. Eriksson, and T. Voigt. *Multi-channel Two-Way Time of Flight Sensor Network Ranging*. In Wireless Sensor Networks - 9th European Conference, EWSN 2012, Trento, Italy, February 15-17, 2012. Proceedings, pages 163–178, 2012.

[10] P. Ferrari, A. Flammini, E. Sisinni, A. Depari, M. Rizzi, R. Exel, and T. Sauter. *Timestamping and Ranging Performance for IEEE 802.15. 4*

CSS Systems. Instrumentation and Measurement, IEEE Transactions on, 63(5):1244–1252, 2014.

[11] E. B. Mazomenos, D. D. Jager, J. S. Reeve, and N. M. White. *A Two-Way Time of Flight Ranging Scheme for Wireless Sensor Networks*. In Wireless Sensor Networks - 8th European Conference, EWSN Germany, February 23-25, 2011. Proceedings, pages 163–178, 2011.

[12] W. Pöttner, S. Schildt, D. Meyer, and L. C. Wolf. *Piggy-Backing Link Quality Measurements to IEEE 802.15.4 Acknowledgements*. In IEEE 8th International Conference on Mobile Adhoc and Sensor Systems, MASS, Spain, October 17-22, 2011, pages 807–812, 2011.

[13] S. Woo, S. Jeong, E. Mok, L. Xia, C. Choi, M. Pyeon, and J. Heo. *Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou* {*MTR*}. Automation in Construction, 20(1):3 – 13, 2011.

[14] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win. *Ranging with ultrawide bandwidth signals in multipath environments*. Proceedings of the IEEE, 97(2):404–426, 2009.

[15] A. Lewandowski and C. Wietfeld. *A comprehensive approach for optimizing ToA-localization in harsh industrial environments*. In Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION, pages 516–525. IEEE, 2010.

[16] J. Prieto, A. Bahillo, S. Mazuelas, P. Fernández, R. M. Lorenzo, and E. J. Abril. *Self-calibration of TOA/distance relationship for wireless localization in harsh environments*. In Communications (ICC), 2012 IEEE International Conference on, pages 571–575. IEEE, 2012.

[17] L. Yu, M. Laaraiedh, S. Avrillon, B. Uguen, J. Keignart, and J. Stephan. *Performance evaluation of threshold-based toa estimation techniques using IR-UWB indoor measurements*. In European Wireless, 2012. EW. 18th European Wireless Conference, pages 1–7. VDE, 2012.

[18] S. Salous, A. Cheema, and X. Raimundo. *Radio channel propagation measurements using a multiband agile chirp sounder*. In General Assembly and Scientific Symposium (URSI GASS), 2014 XXXIth URSI, pages 1–4, Aug 2014.

[19] A. M. Ladd, K. E. Bekris, A. Rudys, D. S. Wallach, and L. E. Kavraki. *On the feasibility of using wireless ethernet for indoor localization*. IEEE Transactions on Robotics, 20(3):555–559, 2004.

[20] J. Rodriguez-Araujo, J. J. Rodriguez-Andina, J. Farina, and M.-Y. Chow. *Field-Programmable System-on-Chip for Localization of UGVs in an Indoor iSpace*. Industrial Informatics, IEEE Transactions on, 10(2):1033–1043, 2014.

[21] *Low cost PCB antenna for 2.4GHz radio: Meander design*, March 2011.

[22] F. Lemi, V. Handziski, A. Wolisz, T. Constambeys, C. Laoudias, S. Adler, S. Schmitt, and Y. Yang. *Experimental Evaluation of RF-based Indoor Localization Algorithms Under RF Interference*. ICL-GNSS 2015, 22-24 June, 2015.

[23] L. Jia and R. J. Radke. *Using time-of-flight measurements for privacy-preserving tracking in a smart room*. Industrial Informatics, IEEE Transactions on, 10(1):689–696, 2014.

[24] Y.-T. Chan, H. Y. C. Hang, and P.-c. Ching. *Exact and approximate maximum likelihood localization algorithms*. Vehicular Technology, IEEE Transactions on, 55(1):10–16, 2006.

[25] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *Automated linear regression tools improve RSSI WSN localization in multipath indoor environment*. EURASIP J. Wireless Comm. and Networking, 2011:38, 2011.

[26] E. Xu, Z. Ding, and S. Dasgupta. *Reduced complexity semidefinite relaxation algorithms for source localization based on time difference of arrival*. Mobile Computing, IEEE Transactions on, 10(9):1276–1282, 2011.

Part II

Comparison of Indoor Localization Solutions

3

Comparability of RF-based Indoor Localization Solutions in Heterogeneous Environments: an Experimental Study

This chapter describes an experimental study that compares multiple indoor localization solutions in heterogeneous environments. Firstly, the presented results in this work show that the performance of localization solutions can be biased if (i) the solution is evaluated (and thus optimized) in one single environment (as demonstrated in Chapter 2) and (ii) if only point accuracy is reported. Additionally, the results of this study confirm the need for standardization of evaluation methodologies. The lessons learned were indispensable for shaping the benchmarking methodology which is elaborated in Part III.

⋆ ⋆ ⋆

**T. Van Haute, E. De Poorter, I. Moerman, F. Lemic,
V. Handziski, A. Wolisz, N. Wirström, T. Voigt**

Published in the International Journal of Ad Hoc and Ubiquitous Computing, Nov. 2014. [IF: 0.900 , Q3]

Abstract The growing popularity of indoor localization research has resulted in a significant amount of research papers describing and evaluating innovative localization solutions. Unfortunately, the results from most of these research papers can not easily be compared since they are evaluated in different environments, use different evaluation criteria and typically tailor their solutions towards a single testbed environment. To evaluate how these different conditions influence the localization performance, in this chapter an exhaustive set of experiments has been set up in which three different localization solutions have been evaluated using multiple metrics in three different test environments: two types of office environments and an industry-like factory environment. None of the used localization solutions was previously optimized for any of these test environments and they were all evaluated under similar conditions, e.g. similar interference conditions and using the same evaluation points. The results reveal several weaknesses in the evaluation methods used in the majority of existing scientific nature of indoor localization solutions. More specifically, it is shown that (i) papers that use self-selected evaluation points can significantly influence their reported accuracy by artificially selecting those evaluation points that outperform other locations, (ii) the reported accuracy can decrease significantly, up to a factor 10, when evaluating an environment for which the localization solution is not specifically tweaked and (iii) many inherent trade-offs between different metrics, such as accuracy, energy consumption and response delay, are hidden by reporting only on the accuracy of the solutions.

3.1 Introduction

3.1.1 Why indoor localization

The emergence of satellite navigation systems - mainly GPS [1] - has resulted in a significant increase of personalized location-based services suitable for guidance, navigation, tracking, recreation, security, etc. However, the use of GPS is limited to outdoor environments, whereas many commercial applications are envisioned in indoor environments. Location-based services are envisioned in many different indoor environments: hospitals, airports, underground mines, detention houses, etc.

A significant amount of work is available in scientific literature describing and evaluating innovative techniques or solutions for localization inside buildings. As a result a wide range of indoor localization solutions has been proposed using a variety of different technologies such as Wi-Fi, infrared, ultrasonic, RF, Bluetooth, 60GHz, etc. However, a major problem is the lack of comparability between indoor localization solutions.

- The majority of evaluations of indoor localization solutions [2, 3] focus mainly on the accuracy of the results whilst ignoring crucial application-level metrics such as scalability, delay, energy consumption, cost, simplicity, etc. Moreover, even the reported accuracy is typically calculated using different calculation methods (average, median, percentiles, etc.), thereby making comparison of solutions is almost impossible.

- In addition, even though each of the targeted application domains has different environmental characteristics, most of existing solutions were evaluated in one specific test environment. As a result, it is impossible to gain insight in the overall performance of these solution under different conditions.

Based on these observations, we argue that a comparative analysis of how a localization approach performs in multiple environments is missing. The main reason for this lack of comparability studies is the significant effort is currently requires to perform localization experiments in multiple experimentation facilities. The main goal of this chapter is to identify to what extent these shortcomings influence the comparability of results in existing scientific literature and to provide suggestions for improvement. Therefore, we implemented three typically used localization approaches and evaluated their performance in multiple test environments using the same evaluation methodology. Amongst the evaluated solutions, we include

- Two popular RF technologies (IEEE 802.11 and IEEE 802.15.4)

- Three localization approaches (ToA, fingerprinting, weighted RSSI)

- Four evaluation metrics (point accuracy, room accuracy, energy consumption, response time)

- Three different test environments: two office environments and an industrial-like open environment.

To the best of our knowledge, we are the first to evaluate multiple localization solutions in multiple environments using the same evaluation procedures.

The remainder of this chapter is organized as follows. Section 3.2 discusses related work, including ongoing efforts to standardize the evaluation of indoor localization solutions. Next, Section 3.3 describes the evaluated localization solutions: (i) a time-of-arrival based IEEE 802.15.4 solution, (ii) a fingerprinting based IEEE 802.11 solution and (iii) an IEEE 802.15.4 RSSI based solution. Section 3.4 discusses the used evaluation methodology and evaluation metrics. Section 3.5 gives an overview of characteristics of the used experimentation testbeds. Afterwards, the localization solutions have been evaluated and the performance results of the solutions are compared and discussed in Section 3.6 for the office environment with brick walls (TWIST), in Section 3.7 for the office environment with plywood walls (w-iLab.t I) and in Section 3.8 for the open industrial like environment (w-iLab.t II). This is followed by a general overview in Section 3.9 where several lessons learned are discussed. Finally, Section 3.10 concludes the chapter.

3.2 Related work

Recently, there has been a growing awareness that a more thorough way of comparing and evaluating localization solutions is needed. This section gives an overview of efforts related to evaluation procedures for indoor localization solutions.

3.2.1 Evaluation procedures for indoor localization

The need for a systematic and objective evaluation methodology has been recognized by several authors [4]. Although no standardized methodologies are currently available, several efforts are being made towards this goal.

- The FP7 EVARILOS project[1] [6] focuses on the **eva**luation of **R**F-based **i**ndoor **lo**calization **s**olutions. The project published a first draft of a benchmarking handbook [7, 8], describing methods to calculate metrics, descriptive methods to describe evaluation environments and methods for deciding which evaluation points to use. The project is also the first to point out that current scientific literature lacks studies on the effect of interference on indoor localization solutions, although interference is expected to be present at most sites where these systems are installed.[2].

[1] The results described in this chapter originate from the project and have first been described in the public EVARILOS deliverable D2.2 “Report on experiments without interference” [5]

[2] The outcome of initial studies on the influence of interference on the localization solutions evaluated in this chapter can be found in EVARILOS deliverable D2.3 “Report on experiments with interference” [9]

- In parallel, ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) have established a joint technical committee, ISO/IEC JTC 1, to jointly work on the ISO/IEC 18305 draft with the aim of standardizing "Test and evaluation of localization and tracking systems" [10][3]. The draft is at the time of writing not yet publicly available, but currently includes a taxonomy of localization solutions and describes a wide range of evaluation scenarios and performance metrics. In contrast to the EVARILOS project, which focuses mainly on RF-based localization solutions, the draft also considers indoor localization solutions that use a wide range of other sensors such as accelerometers.

In contrast to this chapter, the EVARILOS handbook and the ISO/IEC 18305 draft do not include evaluations of different localization solutions in multiple environments, nor do they discuss the inherent trade-offs that are present in current localization solutions.

3.2.2 Evaluation metrics for indoor localization

In more recent surveys, the importance of multiple metrics becomes visible.

- Hui Lui et al. states in [4] that comprehensive performance comparison requires not only accuracy includes but also needs to include precision, complexity, scalability, robustness and cost.
- In the EvAAL project ("Evaluating AAL Systems through Competitive Benchmarking") [11], a competition is held that aims at establishing benchmarks and evaluation metrics for comparing Ambient Assisted Living solutions. For this competition, besides accuracy, also usability metrics are defined such as installation complexity, user acceptance, availability and interoperability with AAL systems [11].
- A significant number of additional metrics can be found in the aforementioned EVARILOS handbook [8] and ISO/IEC 18305 draft [10], both including additional functional metrics, such as response delays, and non-functional (deployment) metrics such as setup time and required infrastructure.

The full list of potential metrics from these sources is very large, especially since many of these metrics can be calculated using multiple statistics

[3] ISO/IEC 18305 is being prepared by Joint Technical Committee ISO/IEC JTC 1, Information technology, Subcommittee SC 31, Automatic identification and data capture techniques, Working Group 5, Real time locating systems. The committee is currently referred to as ISO/IEC JTC1/SC31/WG5.

(percentiles, averages, median, distributions, etc.). Some metrics are important mainly from a theoretical point of view and as such are well-suited for analyzing and improving algorithms of researchers [12], whereas other focus on the performance of end-systems and as such are more important for the industry. Unfortunately, although the above sources strongly emphasize the need for utilizing multiple criteria for evaluating indoor localization solutions, none of these sources mention which of the metrics are considered most important for different application domains, nor do they offer insight on the relation between different metrics (e.g. inherent trade-offs).

Therefore, in the evaluation section of this chapter, we have included four functional metrics that are relevant for most industry deployments and evaluated how these metrics differ in multiple testbed environments: point accuracy, room accuracy, response delay and energy consumption.

3.2.3 Evaluation environments for indoor localization

It is a well-known fact that environmental conditions significantly influence propagation characteristics. Table 3.1 gives an overview of a number of recent research papers evaluating localization solutions and describes the environments they have been evaluated in.

It is clear from Table 3.1 that often indoor localization solutions have been evaluated in office environments, since these are the buildings which are most readily available for researchers. Due to the time-consuming nature of performing localization experiments, most localization solutions are evaluated only in a single environment. However, as will be shown in Section 3.5, office environments can have very different characteristics. Based on existing literature, it is not clear how these differences in environment influence the reported accuracy results. Therefore, this chapter will analyze the performance of multiple localization solutions in three different environments: an office environment with brick walls, an office environment with plywooden walls and an industrial-like open environment.

3.2.4 Evaluation points for indoor localization

In terms of which points to use in an environment to evaluate the performance of a localization solution, two main approaches are possible. For industry-related testing, an evaluation track can be created that mimics typical operations in a building. For example, the path of a person can be recreated and only evaluation points on this path can be used [12]. For more generic, application-independent testing, ideally the evaluation

Table 3.1: Overview of a few existing indoor localization solutions with the related environment, testbed and metrics.

Name solution	**Environment**	**Used Testbed**	**Used Metrics**
Energy efficient solution [12]	Office	Building on campus	Point accuracy & energy consumption
GSM fingerprinting [13]	Office / Home	University, Research Lab, House	Point accuracy
Wi-Fi Bayesian [14]	Office	Their own hallway	Point accuracy
EZ localization [15]	Office	Office floor, Call Center	Point accuracy
Smartphone localization [16]	University	Berkeley Campus	Point accuracy
Wi-Fi in tunnel [17]	Mining	Tunnel in Guangzhou MTR	Point accuracy
Fingerprinting [18]	University	Fourth floor of university building	Point accuracy
UWB fingerprinting [19]	Office / testroom	Anechoic Chamber, Office floor	Number of multipath components

points should be randomly chosen. Unfortunately, most research papers manually select a number of evaluation points based on subjective criteria such as accessibility. As will be shown in Section 3.6, the accuracy of localization solutions can depend strongly depending on the used evaluation points, e.g. points near a wall versus open spaces. As a result, the performance of localization solutions can artificially be 'improved' by selecting mostly evaluation points which perform well for the evaluated solution. It also means that localization solutions that have been evaluated on the same testbed using different evaluation points can not objectively be compared with each other.

As such, it is clear that future evaluations of indoor localization solutions should use standardized evaluation methods. To remedy this, future benchmarking methodologies such as EVARILOS and ISO/IEC 18305 are creating standardized methods for generating evaluation points. For this chapter, all evaluated localization solutions use the same evaluation points in each testbed.

3.3 Evaluated localization solutions

To evaluate how different test environments influence typical localization solutions, we selected three localization solutions that use different wireless technologies and that use different processing approaches for estimating positions. The following localization solutions were selected and implemented.

- An IEEE 802.15.4 based time-of-arrival solution.
- An IEEE 802.11 based fingerprinting solution.
- An IEEE 802.15.4 based RSSI triangulation solution.

Although more accurate solutions exist, these solutions represent the most popular RF-based technologies described in literature.

3.3.1 Particle Filter using ToA and RSSI Measurements

The first solution is designed by N. Wirström et al. [20]. The basic concept behind this localization solution is the following: measurements are performed by letting a stationary node transmit packets to the testbed nodes that reply with a hardware ACK (acknowledgement). The initiating node measures both the time between the transmission of the packet and the reception of the ACK, and stores the RSSI values associated with the

ACK. These measurements are then processed using *Spray*, a particle filter based platform.

The basic idea of the ToF ranging is to estimate the distance between two nodes by measuring the propagation time that is linearly correlated to the distance between the nodes when they are in LoS. Two-way ToF ranging, as opposed to one-way, does not require tight time synchronization between sender and receiver. This is an advantage since tight time synchronization is hard to achieve in wireless sensor networks (WSNs) [21].

The distance between nodes can be calculated according to Equation 3.1 where c is the speed of light, t_{ToF} is the round-trip-time measurements, and t_{off} is an offset time accounting for all processing delays in the system. This includes the time for the sender to transmit the packet, the time the receiver needs to process it, and send the acknowledgement.

$$d = \frac{c}{2}(t_{ToF} - t_{off}) \tag{3.1}$$

The measurements t_{ToF} are computed as $t_{ToF} = \frac{n_{cycles}}{f_{timer}}$, were n_{cycles} is the number of measured clock ticks, and f_{timer} is the frequency of the radio's internal crystal oscillator. In this case $f_{timer} = 12$ MHz. A single measurement is not sufficient, however. The resolution of a single clock allows for a spatial precision equal to $\Delta_d = \frac{c}{2f_{timer}}$. For a 12 MHz clock, the resulting spatial resolution is 12.5 m. To achieve higher resolution, one can average over a series of measurements, as proposed by Mazomenos et al. [22]. This way, sub-clock precision can be achieved.

3.3.1.1 Range Computation Methods

Once the range measurements are collected, they have to be transformed into actual distance measurements. For this, a wide range of computation methods are available. They applied five different methods to the measurements. Four of these use ToF measurements as input, and one use RSSI measurements. The following subsections describe the methods.

Mazo This model builds directly on Equation 3.1. This is the model used by Mazomenos et al. [22]. The calibration step consists of estimating the constant offset t_{off} by averaging over various ToF measurements according to equation 3.2.

$$\hat{t}_{off} = \frac{1}{N}\sum_{i=1}^{N} t_{ToF,i} - \frac{2d_i}{c} \tag{3.2}$$

k-sigma This method was proposed by Pettinato et al. [20]. It uses the variance between measurements taken on different channels to improve range estimations. The idea is that when two nodes are in line-of-sight, most packets will travel the shortest path between the nodes, regardless of the channel being used.

If the two nodes are not in the LoS, however, the different frequencies of the different channels will cause slightly different propagation paths, and result in different ToF measurement values. The concept is captured in Equation 3.3, where σ is the inter-channel standard deviation. Calibration consists of estimating t_{off} and k using linear regression.

$$d = \frac{t_{ToF}}{2} - t_{off} - k\sigma \tag{3.3}$$

Least Squares For this method, the calibration phase consists simply of fitting data to the Equation 3.4, where a and b are estimated using linear regression. This method is model-free in the sense that it does not rely on a physical model.

$$d = a + bt_{ToF} \tag{3.4}$$

Free Space RSSI This method uses the free space propagation model in the Equation 3.5, to transform RSSI measurements to range estimations. In the equation, P_r and P_t are the received and transmitted power, respectively. G_r and G_t are the receivers and the transmitters antenna gains, respectively. λ is the wavelength and L is called the system loss factor.

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{3.5}$$

However, instead of determine these constants individually we combine them into on single constant K as in Equation 3.6, and estimate K using least squares approximation.

$$P_r = K\frac{1}{d^2} \tag{3.6}$$

3.3.1.2 Using Spray to Estimate Location

Once the raw range measurements are transformed to distance estimations, the final location estimations are obtained from *Spray* [23], a particle filter based localization system that can be used to fuse multiple types of measurements simultaneously.

In this case, i.e. using a single range based modality, *Spray* generates *particles* that have both a position and a weight, in a ring-shaped cloud (an annulus) around each testbed node that has an associated range measurement to the node that is to be localized. The distance between each particle and its associated testbed node, is determined by the sum of the range measurement and a zero-mean normally distributed random variable with a given variance.

Each particle is then evaluated using measurements from all the other testbed nodes, on the basis of how the particle's position fits their measurements. This is done by assigning a weight between 0 and 1 to the particle. The more coherent the particle's position is with the measurement, the higher the weight. A final weight for each particle is then computed by multiplying the weights assigned in the evaluation phase.

3.3.2 Fingerprinting Based Localization using Wi-Fi Beacon Packets' RSSI Measurements

Another solution is provided by F. Lemic [24] which is based on fingerprinting. Fingerprinting methods in the indoor localization are generally divided in two phases. The first phase is called the training or *offline* phase. In this phase, the localization area is divided in a certain number of cells. Each cell is scanned a certain number of times for different signal properties, and using a methodology for processing the received data a fingerprint of each cell is created. By using the obtained training fingerprints the training database is created and stored on the localization server. In the second phase, known as the runtime or *online* phase, a number of scans of the environment are created using the user's device. From the scanned data, using the same predefined data processing methodology, the runtime fingerprint is created and sent to the localization server. At the server's side the runtime fingerprint is compared with the training dataset using the matching method. The training fingerprint with the most similarities to the runtime fingerprint is reported as the estimated position. In the section below a general notion of the Wi-Fi fingerprinting is given using beacon packets' RSSI values.

Let K_t and M be respectively the number of Wi-Fi APs used for a localization procedure and the number of training points in a given localization area. Furthermore, let N_t be the number of scans of the area taken at a training point m ($m \in 1, ..., M$). During each scan the vector of RSSI measurements from each visible AP used for localization is collected. This vector has at most K_t elements, but it is possible that it will have less elements if the user's device is not in the range of a number of APs or because beacon packets are lost due to interference. After collecting N_t measurement vectors from different APs at training point i the training matrix $\boldsymbol{S}_i^t$ is created. The matrix $\boldsymbol{S}_i^t$ has K_t rows and N_t columns ($\boldsymbol{S}^t_{K_t \times N_t}$). The matrix of the training measurements from each training cell is preprocessed training data. Based on the methodology that each localization algorithm uses for creating the fingerprint, from the matrices $\boldsymbol{S}^t$ M training fingerprints are created.

A similar procedure, with different parameters, is used for creating the runtime scan of the RSSI measurements. Let K_r be the number of Wi-Fi AP used in the localization procedure and visible to the user's device at a given localization. The number of measurements taken by the user's device is equal to N_r. A runtime fingerprint is a matrix of RSSI values $\boldsymbol{S}^r_{K_r \times N_r}$. The fingerprint is created using a methodology defined in the fingerprinting based localization algorithm.

The purpose of fingerprint based localization algorithms is to accurately detect the similarities between training dataset and run-time fingerprints. Due to the time and energy constrains of a (usually wireless) user's device, the number of measurements in the runtime fingerprint N_r is usually smaller than the number of measurements taken while collecting training fingerprints N_t. For this reason, the number of measurements given as an input to a localization algorithm is equal to N_r. Furthermore, only a subset of RSSI measurements from the APs that are common to both training and runtime fingerprint is given to the second phase of the localization algorithm.

For the evaluation, they used three fingerprint based indoor localization algorithms which have been used in previous research work.

Distance of RSSI Confidence Intervals The Weighted Sum (WS) of the RSSI Confidence interval approach in the fingerprint-based indoor localization uses the vector of confidence intervals in both training and running phase of the localization to estimate a user's position. Each confidence interval is generated using the RSSI values received by corresponding AP. Let the confidence interval from the access point *i* during the training phase be [T_i^-, T_i^+]. Furthermore, let the confidence interval from the access point *i* during the runtime phase be [R_i^-, R_i^+]. The fingerprint of the cell (in training or runtime phase) is a vector of the given confidence intervals for all APs used in localization procedure. From here, it is possible to define the *weight* between the running confidence interval and each cell in the training confidence interval. The weight between the training point *t* and the running point is given by:

$$w(t) = \begin{cases} \frac{T_i^+(t)-R_i^-}{R_i^+-T_i^-(t)} & if\ (T_i^-(t) < R_i^- < T_i^+(t) < R_i^+) \\ \frac{R_i^+-T_i^-(t)}{T_i^+(t)-R_i^-} & if\ (R_i^- < T_i^-(t) < R_i^+ < T_i^+(t)) \\ 1 & if\ (T_i^-(t) \leq R_i^- < R_i^+ \leq T_i^+(t)) or \\ & (R_i^- \leq T_i^-(t) < T_i^+(t) \leq R_i^+) \\ 0 & if\ (T_i^-(t) < T_i^+(t) \leq R_i^- < R_i^+) or \\ & (R_i^- < R_i^+ \leq T_i^-(t) < T_i^+(t)) \end{cases}$$

Adding all weights will compute the weighted sum, i.e. the WS distance. The computed distance indicates the similarity between the cell in the training dataset and the runtime fingerprint. The cell with the maximum weight in the WS distance of confidence intervals approach is considered the estimated position.

ED of Averaged RSSI Vectors The Euclidean Distance (ED) of the averaged RSSI vectors is one of the most basic and well known algorithms used for fingerprint based indoor localization algorithms [25]. The input to the matching method is an average value of RSSI measurements obtained from each AP used for localization in both training and runtime phase, where $K_{r,t}$ is the length of the vector. Let $\boldsymbol{\mu}_{t,m} = [\overline{RSSI}_{t,1}, ..., \overline{RSSI}_{t,k}, ..., \overline{RSSI}_{t,K_{r,t}}]$ be the vector of averaged RSSI values from each AP obtained during the training phase at cell $m \in 1, ..., M_t$, i.e. the training fingerprint. In the same manner, let $\boldsymbol{\mu}_r = [\overline{RSSI}_{r,1}, ..., \overline{RSSI}_{r,k}, ..., \overline{RSSI}_{r,K_r}]$ be the vector of averaged RSSI values from each AP obtained during the runtime phase, i.e. the runtime fingerprint. The distance between the training fingerprint at the cell m and the runtime fingerprint is given as:

$$D_E(\boldsymbol{\mu}_{t,m}, \boldsymbol{\mu}_r) = |\overline{\boldsymbol{\mu}}_{t,i} - \overline{\boldsymbol{\mu}}_{r,i}| \tag{3.7}$$

The distance $D_{EU}(\boldsymbol{\mu}_{t,m}, \boldsymbol{\mu}_r)$ is the ED distance between the vectors of averaged RSSI values of the cell m and runtime point. The cell with the smallest distance (also called smallest weight) is reported as the estimated position.

KL Distance of MvG Distributions of RSSIs The third fingerprinting based indoor localization algorithm uses the Kullback-Leibler (KL) distance between the Multivariate Gaussian distributions of RSSI measurements from each AP used in the localization procedure [25]. The algorithm assumes that the RSSI values from each AP are distributed according to the Multivariate Gaussian distribution. In other words, the distribution of the

RSSI values from each AP at one cell can be written as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In the same manner as in the previously presented algorithm, let $\boldsymbol{\mu}_{t,m}$ and $\boldsymbol{\mu}_r$ be the vectors of the averaged RSSI values from each AP in training phase at the cell m and in the running phase, respectively. Furthermore, let the $\boldsymbol{\Sigma}_{t,m}$ and $\boldsymbol{\Sigma}_r$ be the covariance matrices of the RSSI measurements at training cell m and running point respectively. The Multivariate Gaussian distributions of the training point m and running point can then be written as $\mathcal{N}_{t,m} = \mathcal{N}(\boldsymbol{\mu}_{t,m}, \boldsymbol{\Sigma}_{t,m})$ and $\mathcal{N}_r = \mathcal{N}(\boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$ respectively.

$$\begin{aligned} D_{KL}(\mathcal{N}_{t,m}, \mathcal{N}_r) = \frac{1}{2}((\boldsymbol{\mu}^S_{i,T} - \boldsymbol{\mu}^S_R)^T (\boldsymbol{\Sigma}^S_{i,T})^{-1} (\boldsymbol{\mu}^S_{i,T} - \boldsymbol{\mu}^S_R) + \\ tr(\boldsymbol{\Sigma}^S_{i,T} (\boldsymbol{\Sigma}^S_{i,T})^{-1} - \boldsymbol{I}) - ln|\boldsymbol{\Sigma}^S_R (\boldsymbol{\Sigma}^S_{i,T})^{-1}|) \end{aligned}$$

where *tr*$(\cdot)$ denotes the trace of a matrix (sum of its diagonal elements) and **I** is the identity matrix. The matching method reports the cell with the smallest KL distance as the estimated position.

PH Distance of RSSI Quantiles Finally, as a fourth fingerprinting method, we propose a new approach using quantiles of the RSSI values from each AP for creating fingerprints and the Pompeiu-Hausdorff (PH) distance for estimating the similarities between the training and runtime fingerprints. Using the quantiles for indoor localization purposes is frequently used in robotics, where robots are using quantiles of images of the environments in order to localize themselves [26]. PH distance is usually used in image processing for pattern recognition and measuring the dissimilarities between shapes. As far as we know, using a combination of quantiles of RSSI distributions and PH distance for location estimation has not been proposed and examined in literature. We find this approach promising because a higher amount of information is provided to the matching method. In other words, in our opinion using only the vector of averaged RSSI values and the covariance between measurements between different APs may not be sufficient for precise localization. In our case the q-quantile of the RSSI measurements from each AP is calculated in two steps. The first one computes the cumulative distribution functions (CDF) of the RSSI measurements from each AP. The second step calculates the quantiles, i.e. the RSSI values with probabilities $k/(q-1)$, where $k = 0, 1, ..., q-1$. The result of the quantile calculation in both training and runtime phase is a quantile matrix $Q_{K,q}$, where K is the number of APs visible at the given location and q is a number of quantiles. The similarities between the RSSI quantiles from the training fingerprints and the runtime fingerprint are computed using the PH distance metric. The PH distance between two sets of quantiles is given as follows:

$$D_{PH}(Q_1, Q_2) = \max_{q_{1,k} \in Q_1} (\min_{q_{2,k} \in Q_2} (d(q_{1,k}, q_{2,k}))) \tag{3.8}$$

where $d(q_{1,k}, q_{2,k})$ is the Euclidean distance (ED) measurement. The training cell with the smallest PH distance is reported as an estimated location.

3.3.3 Hybrid Model: Proximity & weighted RSSI

A final localization solution [27] that has been implemented and evaluated is a hybrid combination of a range-based and a range-free algorithm. It includes a range-based location estimator based on weighted RSSI values. The main idea of RSSI is that the transmission power P_T directly affects the received power P_R of a signal. Using the Friis transmission equation, the linear relationship can be stated as follows.

$$P_R = P_T * G_T * G_R \left(\frac{\lambda}{4\pi d} \right)^2 \tag{3.9}$$

In the equation G_T, G_R are the gains of transmitter and receiver, respectively. λ is the wavelength of the signal and d is the distance between sender and receiver. The RSSI can be defined as the ratio of the received power to the reference power P_{Ref}.

$$RSSI = 10 * \log \frac{P_R}{P_{Ref}} \tag{3.10}$$

Each RSSI value can be matched with a certain distance. The proposed algorithm in [27] not only uses the RSSI values to measure the distance between a fixed and mobile node, but also the distance between the fixed nodes. These values function as weight factors for the distance calculation between the fixed and mobile node. These weight factors are shown in Figure 3.1 as w_{12}, w_{13} and w_{23}. The distance from M to, for example, B_1 can be calculated as follows:

$$Distance(M, B_1) = \frac{RSSI(M, B_1) * w_{12} + RSSI(M, B_1) * w_{13}}{2} \tag{3.11}$$

whereby w_{ij}:

$$w_{ij} = \frac{Dist(B_i, B_j)}{RSSI(B_i, B_j)} \tag{3.12}$$

Previous results [28] prove that these weight factors add value to the accuracy. A drawback of the RSSI technique is that these measurements are very sensitive to the environment and any changes in it. The relationship between the distance and RSSI is room dependent. For example, signals in a long corridor propagate much further because they reverberate through the long walls.

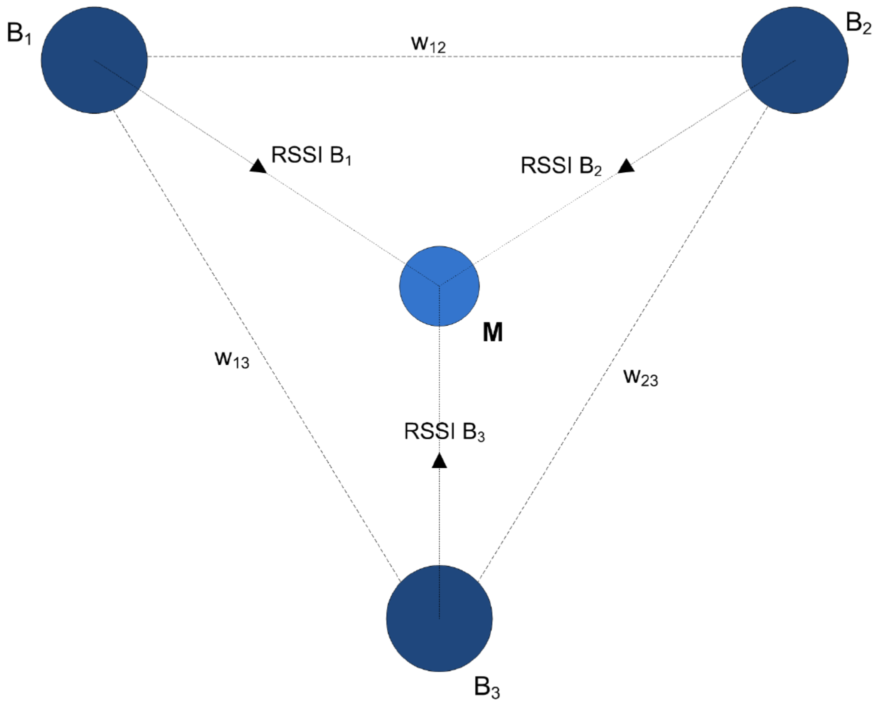


Figure 3.1: The graphical scheme of the Weighted algorithm. Circles B_1, B_2 and B_3 represent the anchors whilst M acts as the mobile node. $RSSI\ B_1$, $RSSI\ B_2$ and $RSSI\ B_3$ are the RSSI values between the mobile and the anchor nodes whilst w_{13}, w_{12} and w_{23} are the RSSI values between the anchor nodes mutual.

In contrast to the technique above, range-free algorithms do not take RSSI-values into account. If a mobile sensor node has a range of 10 meters, then a fixed node can only receive his messages if the mobile node is maximum 10 meters away. This is the only information that is used to calculate the position of a mobile node. For this approach, it is important that the transmission power is well configured. If the power is too low, the mobile node could be out of range between two fixed nodes. On the other hand, if the power is too high, too many fixed nodes will receive the beacon and a wrong estimation could be made.

The latter problem can be solved by using a centroid algorithm. This is only useful if there is a set of fixed nodes with an overlapping coverage area. The beacon of the mobile node is received by multiple fixed nodes. In order to determine the position, the centroid of all the receiving fixed nodes is calculated:

$$\begin{aligned} x_M &= \frac{\sum_{n=0}^{k} x_n}{k} \\ y_M &= \frac{\sum_{n=0}^{k} y_n}{k} \end{aligned} \tag{3.13}$$

In theory, this algorithm would give a 100% guarantee that room-accuracy is possible. However, experiments have shown that this is not always the case. If the walls are small enough and/or do not strongly attenuate the signal, signals can go through and a fixed node in a different room can receive up the beacon. To prevent incorrect location estimation, extra logic can be added to the algorithm.

The extra logic takes the form of additional environmental metadata. Suppose we have the exact coordinates of all the walls, doors and nodes inside a building. Knowing that every beacon has an index number, the direct path could be checked between the two fixed nodes who received the consecutive beacons. If the mobile node goes from one room to another, without using a door, then the last beacon can be dismissed. For example (Fig. 3.2) when node A_2 receives a beacon and the next beacon is received by node B_2. It is impossible to move directly from A_2 to B_2 without passing nodes A_1 and B_1. So the message that was received by beacon B_2 will be rejected.

With this optimization room-accuracy can be guaranteed. Still, this solution has the drawback that a lot of fixed infrastructure sensor nodes are necessary to retrieve good results. If the network is sparse distributed, then the algorithm would not work properly.

Finally, for the evaluation of this solution, experiments were performed using four different TX power levels (TX3, TX7, TX19 and TX31), as shown in Table 3.2.

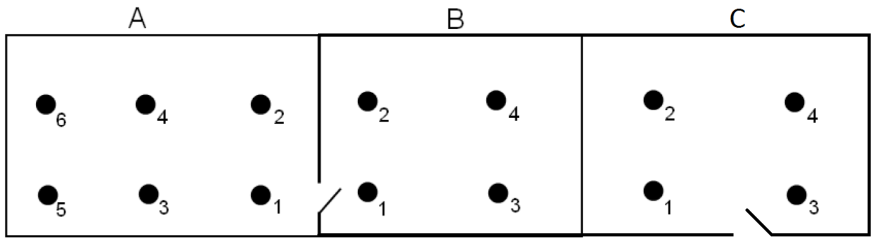


Figure 3.2: Three neighbouring offices

Table 3.2: Used TX power levels for the weighted RSSI localization experiments

Tx power level	Output power [dBm]
3	-25
7	-15
19	-5
31	0

3.4 The Benchmark methodology

3.4.1 Introduction

One of the major problems of indoor localization is the challenge of reproducing research results in real life scenarios and the inability to compare their performance due to evaluation under individual, not comparable and not repeatable conditions. Therefore, contrary to previous approaches, our benchmarking approach does not focus exclusively on the accuracy of the evaluated localization approach, but also considers other performance measures that are relevant from the point of view of practical deployment of localization solutions such as energy efficiency and response time. Due to variation in the sensibility of different use-case scenarios on the individual metrics, the methodology cleanly decouples between evaluating individual metrics and calculation of a final score used for ranking. As illustrated on Figure 3.3, after collecting a set of measurements necessary for the calculation of the individual metrics, the methodology envisions the use of weighting factors and thresholding for the calculation of the final ranking score, reflecting the different impact of the individual metrics for the particular application scenario of interest.

3.4.2 Used metrics

The metrics that will be used for the evaluation of the solutions will have a critical impact on the final score. A classical mistake by other comparison

Figure 3.3: Transform measurements to scores using metrics

and evaluation tools is only using the point accuracy as a reference for a good or bad working solution.
In this chapter, we will take others metrics into account as well, which are defined in the EVARILOS benchmarking handbook [7].

3.4.3 Used scenarios

Each solution is evaluated using a predefined scenario in each testbed. These are based on the generic scenario descriptions of the EVARILOS benchmarking handbook. In the next paragraphs, we will describe each scenario of each testbed. A detailed overview of each testbed is given in Section 3.5.

3.4.3.1 TWIST

The scenario is instantiated on the 2^{nd} floor of the TWIST testbed, and can be characterized as a “small office environment” according to the EVARILOS benchmarking handbook. The evaluation points used to evaluate the localization solutions are shown in the Figure 3.4.

3.4.3.2 w-iLab.t I

At w-iLab.t I, we will use the third floor to execute the experiments. On this floor, 57 nodes are available for the experiments. An overview of the third floor is given in Figure 3.11. There is no actual difference between the green and the blue dots, it is for reservation purposes only.

Unfortunately, not the whole floor can be considered as test area. Some private offices, technical staff room etc. are not available for measuring. The unreachable zones are marked with a red layer. In order to define the measurement points, we used a grid (see Figure 3.5). The decision has to be taken without premeditation. Therefore, a randomizer is used. To avoid measurement points close to each other, making an unbalanced distribution, the principle of the Latin Square is applied.

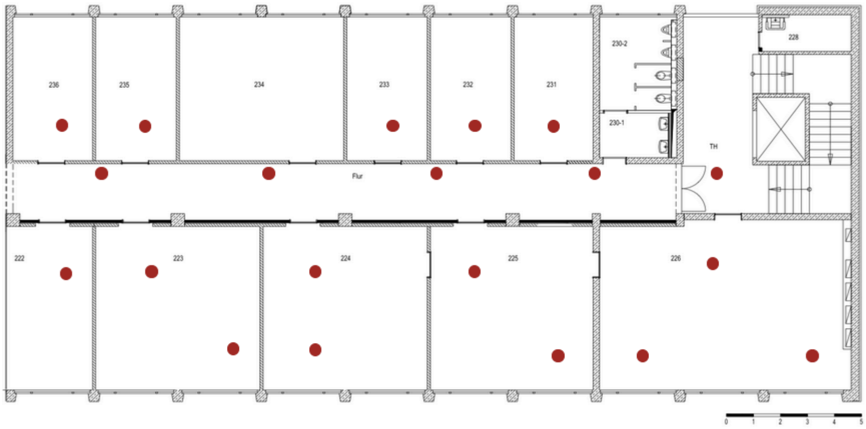


Figure 3.4: TWIST evaluation points utilized for the first benchmarking scenario

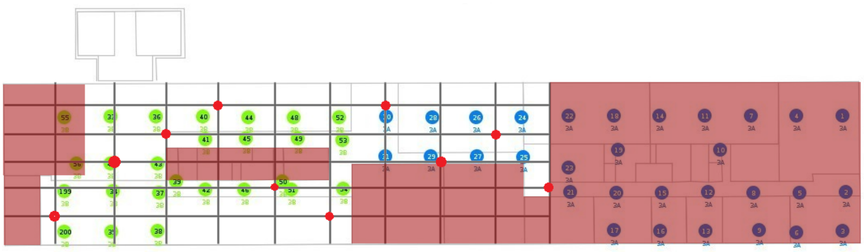


Figure 3.5: Measurement points in the w-iLab.t I office environment

3.4.3.3 w-iLab.t II

In this testbed, the 26 measurement points are well spread over the area. This is shown in Figure 3.6. Special in this setup is that there is no physical person present in the building. Everything is controlled remotely using robots. In this way, the repeatability of the measurement point is very high. On the other hand, this “open environment” is made of metal walls and contains a lot of metal objects, making it very challenging for accurate localization.

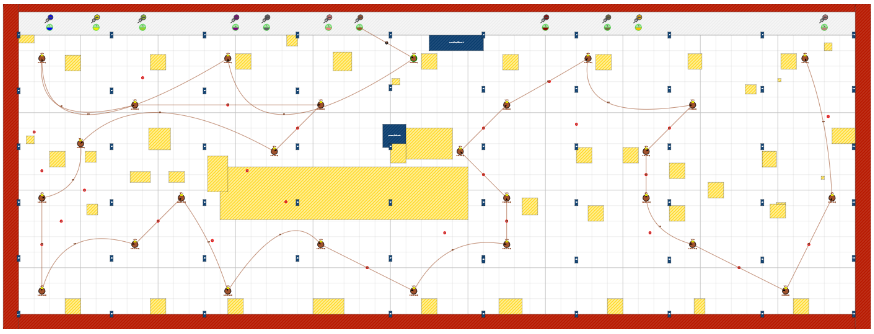


Figure 3.6: The measurement points of w-iLab.t II in Zwijnaarde

3.5 Test environments

3.5.1 TWIST in Berlin

The TKN Wireless Indoor Sensor Network Testbed (TWIST) is a multiplatform, hierarchical testbed architecture developed at the Technische Universität Berlin. The TWIST instance at the TKN office building is one of the largest remotely accessible testbeds. It has 204 SUT sockets, currently populated with 102 eyesIFX and 102 Tmote Sky nodes (Figure 3.7). The nodes are deployed in a 3D grid spanning 3 floors of an office building at the TUB campus, resulting in more than 1500 m^2 of instrumented office space. In small rooms ($\sim$14 m^2), two nodes of each platform are deployed, while the larger ones ($\sim$28 m^2) have four nodes (Figure 3.8). This setup results in a fairly regular grid deployment pattern with intra node distance of 3m. Within the rooms the sensor nodes are attached to the ceiling.

In addition to the described sensor network, the TWIST infrastructure consists of multiple Wi-Fi access points. Deployed Wi-Fi access points are commercial of-the-shelf TL-WDR4300 routers (Figure 3.9). The Wi-Fi routers can serve two functions. They can be used as a part of the localization solution, if particular solution requires Wi-Fi anchor points. In

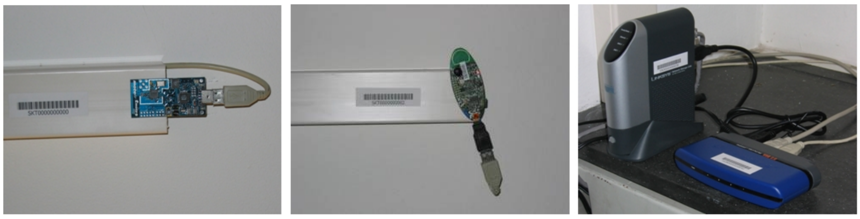


Figure 3.7: TWIST testbed: Nodes

Figure 3.8: TWIST testbed: Map (2^{nd} floor)

the same time, some routers can also be used for creating different types and amounts of IEEE 802.11 traffic in order to generate controlled Wi-Fi interference.

Figure 3.9: TWIST testbed: Access Points

For supporting mobility and automation of the localization measurements multiple TWISTbot robotic platforms (based on the TurtleBot design from Willow Garage), are used. Their function is to carry nodes that need to be localized through the measurement points and report the ground truth position.

Furthermore, the TWIST infrastructure is complemented by several WiSpy sensing devices: these are low-cost spectrum scanners that monitor activity in the 868 MHz, 2.4 and 5 GHz spectrum, and output the measured RF energy and the quality of the received signals. Also, for more precise sensing of the wireless environment spectrum analyzers are used. Finally, except for the before mentioned Wi-Fi routers and TWIST platform, for generating interference Rohde & Schwarz signal generator is used. Signal generator can be used for generating arbitrary signals, and it will usually be used for simulating microwave interference.

3.5.2 w-iLab.t I at De Zuiderpoort

The w-iLab.t I testbed is located at “De Zuiderpoort” in Ghent, Belgium. The infrastructure is distributed on three floors (18x90m) of the iMinds office (Figure 3.11). The network consists of 200 nodes. Every w-iLab.t node is generic and is equipped with one or more sensor nodes, an intermediate node with 2 Wi-Fi 802.11 radios, the environment emulator and a Bluetooth interface.

The sensor nodes are Tmote Sky motes that use a Tiny OS development environment. They consist of a TI MSP430 processor running at 8MHz, 10KB of RAM, 1Mbit of Flash memory and an IEEE 802.15.4 compliant Chipcon CC2420 radio operating at 2.4GHz with a maximum indoor range of approximately 100 meters. Each node includes sensors for light, temperature, and humidity. The hardware setup is extendable with a large variety of other radios (e.g Software Defined Radio, sensing engine), as long as the radio has a USB or RS232 serial interface.

The intermediate nodes (called iNodes, Figure 3.10) are Alix 3C3 devices running Linux. These are mini PC's equipped with Ethernet, USB, serial, vga, audio and two IEEE 802.11 a/b/g interfaces. All the iNodes are connected to the management backbone using Power-over-Ethernet switches, making it possible to power up/down the iNodes as needed without physical interaction with the iNodes. The iNodes can become an active member of the experiment as it is possible to adjust the kernel, the driver, to add click router code or to add java-based applications.

Finally, the Environment Emulator (EE) is located in between the iNode and the sensor node. Using the EE, it is possible to emulate the behavior of any type of sensor or actuator without the need for real sensor/actuator hardware or the development of a full-blown sensor application. It is possible to emulate the battery depletion, depending on the real life power consumption of the sensor node. When the node's battery is depleted or the node is destroyed (e.g., in an explosion), the node can be switched off. The EE can be programmed to emulate a sensor event (e.g., temperature rise, motion detection), an actuator event or to support voice streams. Further, the EE can be used to monitor the energy consumption of each individual sensor. Altogether, this means that it is possible to assess the complete usability of a certain wireless sensor and actuator network application or protocol in a real-life environment. The initial core of w-iLab.t was based on the widely used MoteLab testbed from Harvard University. This building belongs to the category “Plywooden walls” and the size is “Big”.

This is a classic office environment where multiple devices communicate wireless with each other. Laptops using Wi-Fi and bluetooth, smartphones using the 3G network. Here we will consider typical office applications like email, file transfer, video/audio conferencing and web surfing). The office environment is a live environment. Meaning the interference in this testbed is uncontrolled. During daytime several people are working in these buildings. So the w-iLab.t is a testbed with very realistic office interference. The sacrifice of this realistic office environment is the uncontrollable interference.

The w-iLab.t testbed is centrally managed for control and monitoring purposes. It supports easy configuration and deployment, including installation of new software, protocols and middleware components via an intuitive web-based interface. Registered users can upload executables, associate those executables with the nodes (both sensor nodes and iNodes) to create a job, and schedule the job to be run on w-iLab.t. During the job all messages and other data are logged to a database, which is presented to the user upon job completion and then can be used for processing and visualization.

All the possibilities of the complete testbed, the environment emulator scenarios and events, a visualization and a graphical analysis tool, are accessible through a web interface. The visualization tool can visualize any type of node status and/or link information on a map of the building, while the graphical analyser plots out the data. The information for both tools is gathered from the database through the use of user customizable MySQL statements, making it extremely flexible. External users can access the testbed over a secured OpenVPN connection.

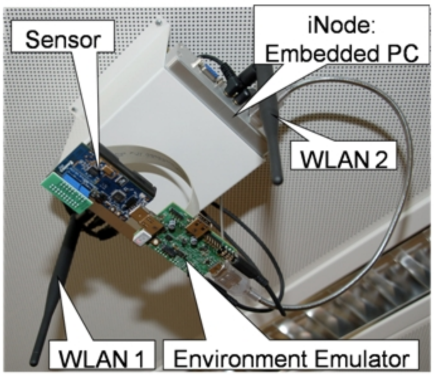


Figure 3.10: iNode mounted to the ceiling of the w-iLab.t I wireless testbed

3.5.3 w-iLab.t II in Zwijnaarde

The w-iLab.t II testbed is located in “Zwijnaarde”, above a cleanroom. At this location, there is (almost) no interference. It is one open space where 60 fixed nodes are distributed over an area of 70m x 25m (Figure 3.12). In this environment there are also 20 mobile nodes. These nodes are based on a vacuum cleaning robot and are extended with a radio for remote control and accurate positioning algorithms (with rasters on the floor). Due to the fact that the movement of these robots is controlled, mobility is reproducible. The fixed nodes are marked with blue spots while the mobile nodes have orange spots on the map in Figure 3.12.

Every node location contains (i) a Zotac embedded PC, (ii) an environment emulator (see w-iLab.t), (iii) an iMinds Rmoni sensor node, (iv) a Bluetooth dongle and some of them have a web cam. These nodes are remotely powered by Racktivity PDU’s.

3.6 Results in TWIST

3.6.1 RSSI and ToA with Particle Filter

In this section we evaluate the particle filter localization approach outlined in Section 3.3.1. Table 3.3 presents the summarized results. In this table,

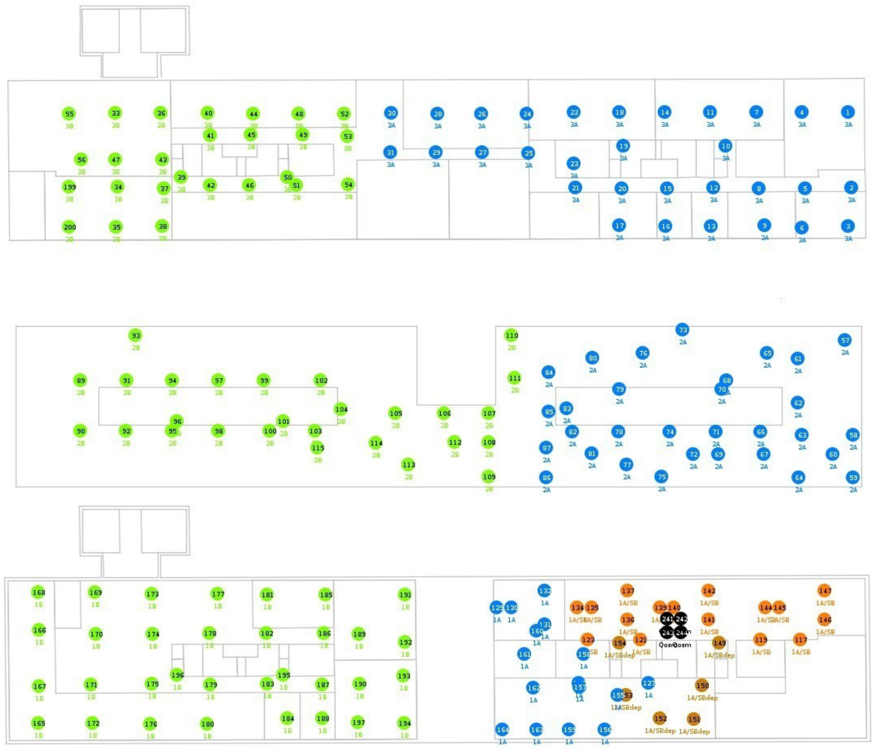


Figure 3.11: The w-iLab.t I wireless testbed: map

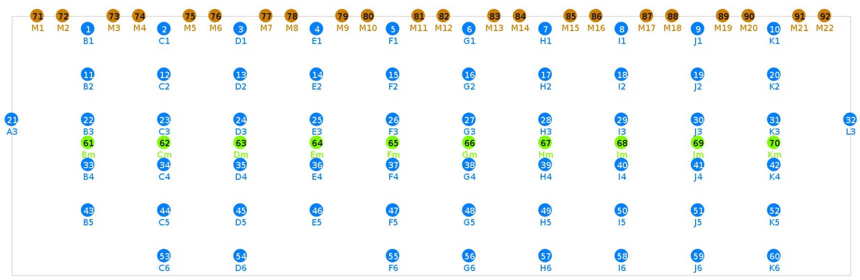


Figure 3.12: The w-iLab.t II wireless testbed: map

“Min.” stands for Minimum, “Max.” for Maximum and “RMS” for Root Mean Square. These abbreviations are also used in the other tables.

Table 3.3: Statistical information about the particle filter algorithms’ performance in TWIST

	RSSI	**ToA**
Average error [m]	4.35	5.56
Min. error [m]	0.62	0.68
Max. error [m]	12.99	22.47
Median error [m]	3.22	3.91
RMS error [m]	5.28	7.11
Room accuracy [%]	45.00	30.00
Response time [ms]	14 285	14 282

Measurements are collected for over a minute at each measurement point. Most of this time is spent trying to reach nodes that are not within reach, and finding the channel that a reachable testbed node currently is using. The response time could be decreased significantly by dedicating a single channel for communication to be used before starting the ranging phase. Moreover, range estimations do not improve significantly after 1300 measurements, as shown in Figure 3.13. Therefore we use only the first 1300 collected values for our range estimations. This also helps limiting the response time because each measurement takes on average 4 ms. We use the same approach for the RSSI measurements, although the figure shows that 50 measurements are likely to be enough. The figure also shows that after approximately 500 measurements the ToA based method performs better.

Figure 3.14 shows the CDFs for the absolute range errors and the localization errors. The RSSI based range estimation performs better than the ToA based estimation, although Figure 3.13 shows that ToA should give better results for a high number of measurements. The reason for this is that only about 50% of the pair-wise ranging procedures result in 500 measurements or more, and only about 10% result in 1300 measurements or more.

The power consumption of both the target node and the testbed nodes is approximately 105 mW. It is computed as the mean of the transmission and reception power consumptions. The energy consumption per node is especially important when battery powered devices will be used, since it directly impacts the lifetime of a battery-powered localization solution. The node energy consumption can also be used to calculate the overall energy consumption. The infrastructure nodes are always on, and a total of 68 testbed nodes are used. As a result, the continuous total power consumption can be calculated to be 7.1 W for the infrastructure. The mobile node

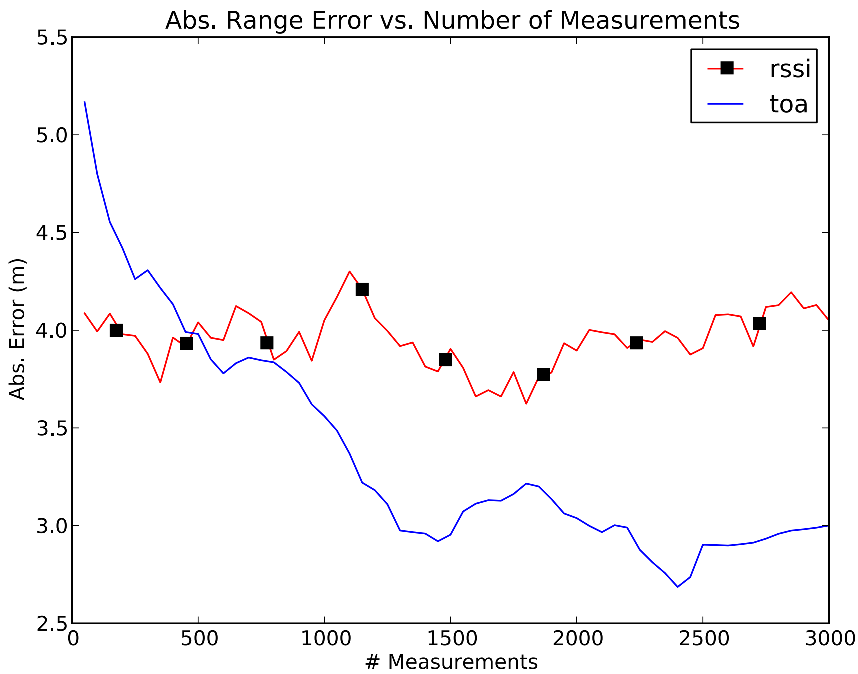


Figure 3.13: The absolute range error for ToA decreases with the number of measurements until approximately 1300 measurements. The RSSI error fluctuates about the same value, and is not improved by additional measurements

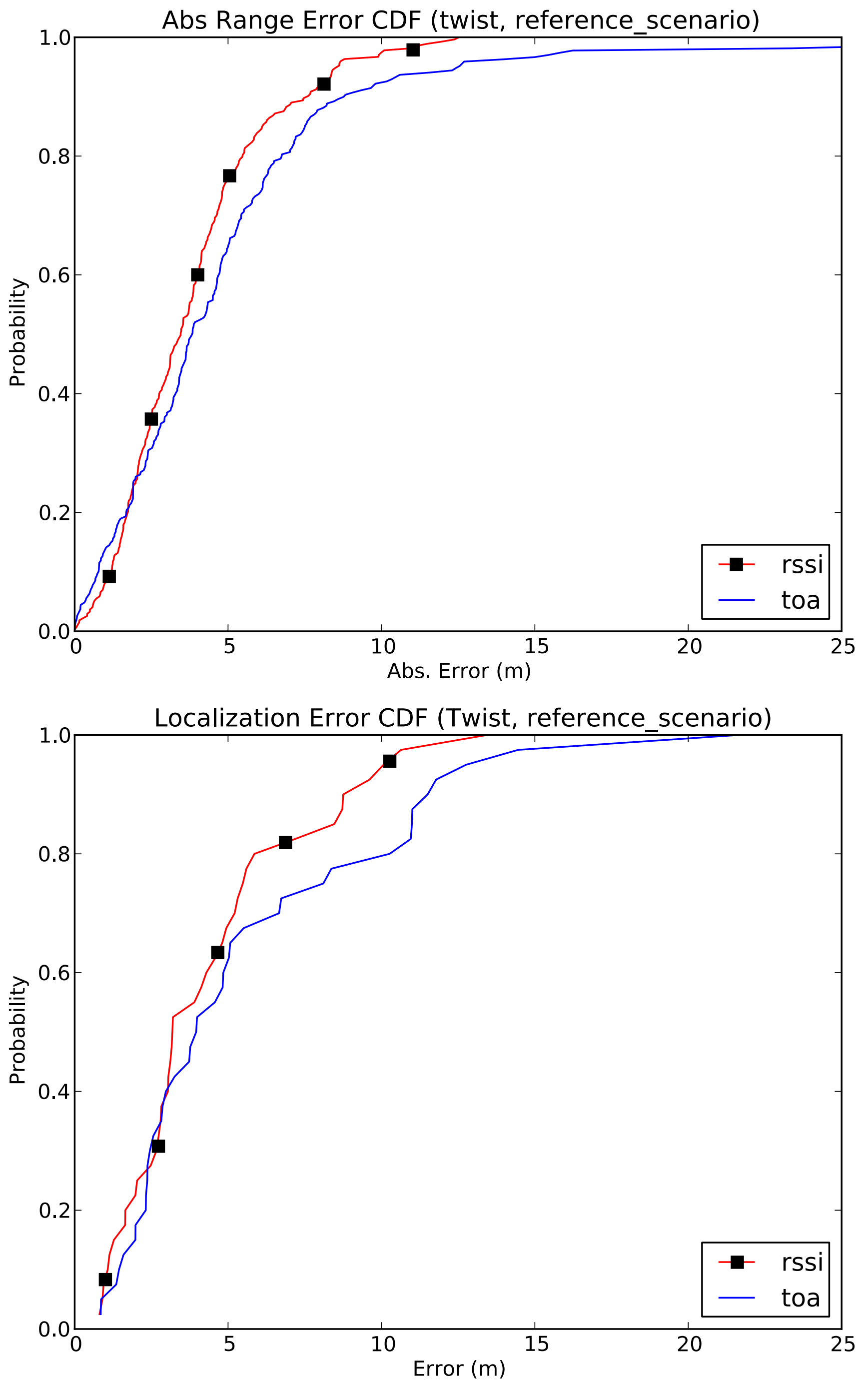


Figure 3.14: CDFs for the absolute range error (top) and the localization error (bottom).

is only on during the response time, which is in the order of 15 s, resulting in an average energy consumption of 1.5 J per measurement.

3.6.2 Fingerprinting

This section evaluates the fingerprinting localization approach described in Section 3.3.2. The accuracy results are shown in Table 3.4. The results show that the PH Distance of RSSI Quantiles give comparable results with the ED Distance of Averaged RSSI Vectors in office scenarios (see also Figure 3.15).

The results also show that, when more beacons are collected and thus the response time increases, the PH has a slightly better overall performance in terms of accuracy. This improvement will be more emphasized in the open space scenario (see results in Section 3.8). The minimum error of all solutions equals zero, which is possible because some of the fingerprints taken during the training set are at the same locations that were used for the evaluation of the algorithm.

Table 3.4: Statistical information about the performance of fingerprinting algorithms in TWIST testbed

	KL	ED	PH
Average error [m]	2.77	2.16	2.02
Min. error [m]	0.00	0.00	0.00
Max. error [m]	5.71	6.35	6.35
Median error [m]	2.98	2.48	2.52
RMS error [m]	3.39	2.95	2.79
Room accuracy [%]	50.00	80.00	85.00
Response time [s]	35.67	35.11	35.12

Since a localization solution, in general, contains many configurable parameters, we expect that they will typically be offered to end users using predetermined configuration setting. As such, it is important to be aware of the inherent trade-offs that are made by the developer of the solution. This is especially important when considering also additional metrics such as the response time. For this solution, the time during which fingerprints are collected (e.g. the time needed before a location estimate could be generated) was set to 35 seconds (excluding the off-line time required for fingerprinting). Figure 3.16 and 3.17 show the trade-offs between response and point and room level accuracies for fingerprinting based solutions in TWIST. As more fingerprints are collected, a better match can be made in order to better estimate the position. Lower response times are possible, at the cost of decreased accuracy. Especially more complex algorithms (such as the PH distance) require more samples to estimate the distributions of the RSSI values. As such, when comparing different localization solutions, the targeted response time has an important influence on the selection of the best algorithm.

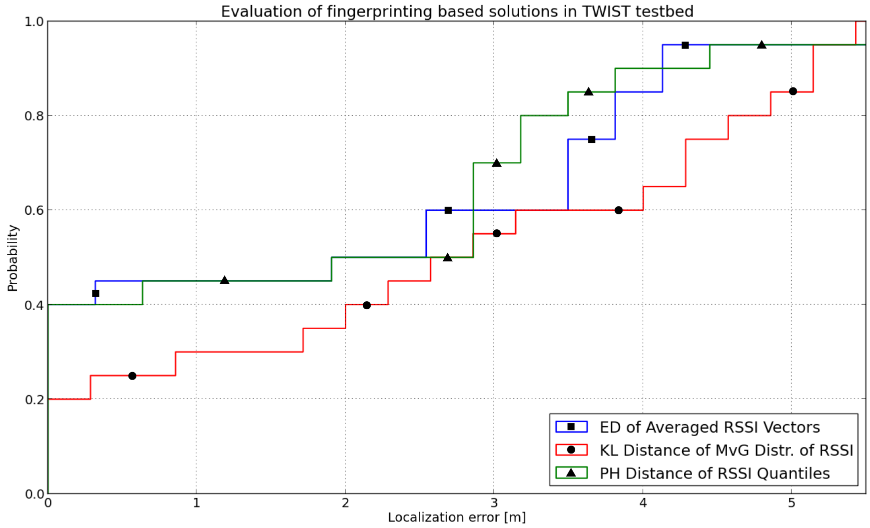


Figure 3.15: CDF of the localization error of fingerprinting based solutions in TWIST testbed

Finally, the energy consumption of the infrastructure nodes (TPLINK 4300 router) is on average 0.5W, whereas the energy consumption of the used mobile devices (MacBook Pro AirPort Extreme NIC) was on average 7W.

3.6.3 Proximity & weighted RSSI

This section evaluates the RSSI based localization approach described in Section 3.3.3. The obtained accuracy is summarized in Table 3.5 for different transmission powers. The average accuracy is relatively low: the concrete walls in the building cause unpredictable signal attenuation, resulting in less accurate estimations of the true location. Using lower transmission powers causes less signals to propagate to multiple rooms, hence the better performance of low transmission powers. A cumulative distribution function of the errors is shown in Figure 3.18.

To estimate the position, the anchor points collect RSSI values from the beacons transmitted by the mobile node. All these RSSI values are collected and merged in the position calculator. There, a translation from RSSI values into coordinates is made. For low transmission powers, the corresponding response delay is about 1.5 seconds (exact values are given

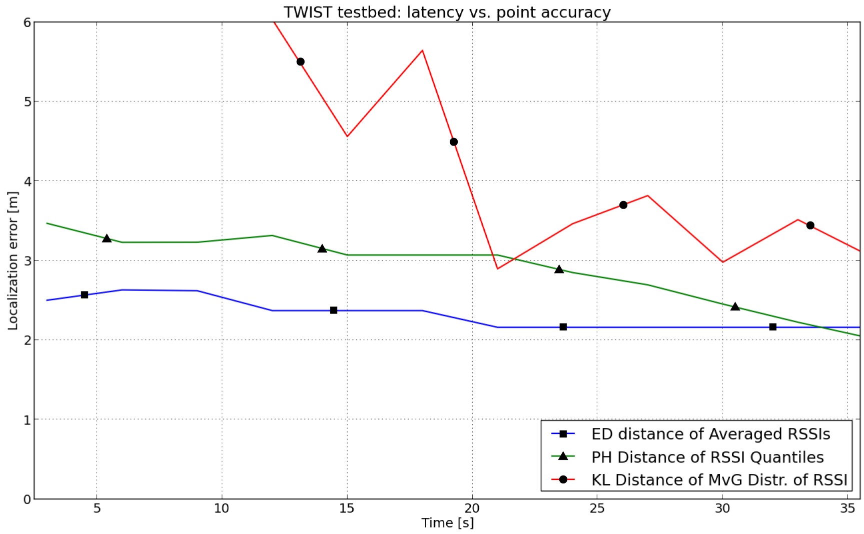


Figure 3.16: Fingerprint collection delay versus point accuracy

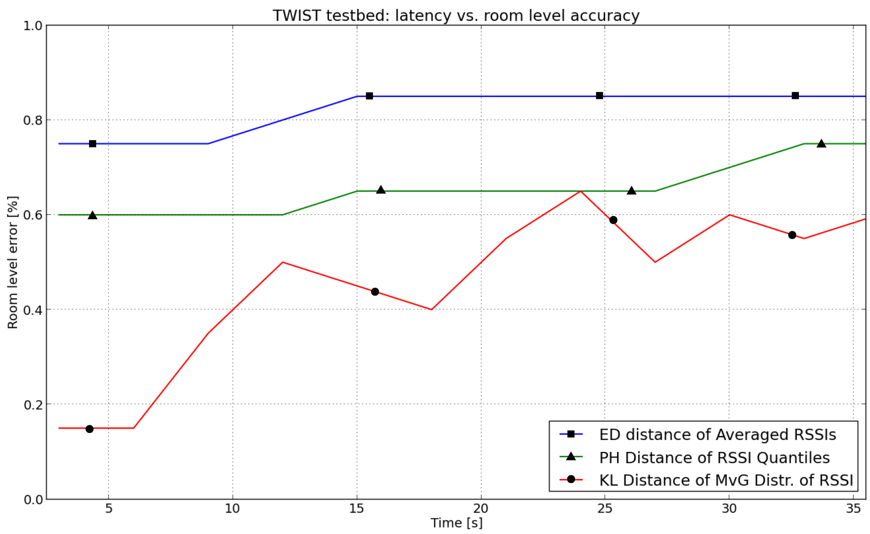


Figure 3.17: Fingerprint collection delay versus room level accuracy

Table 3.5: Statistical information about the hybrid algorithms' performance in TWIST

	Tx3	Tx7	Tx19	Tx31
Average error [m]	4.63	7.08	6.93	8.31
Min. error [m]	0.75	0.83	0.80	0.82
Max. error [m]	10.20	17.52	18.93	19.31
Median error [m]	4.39	6.81	6.68	8.63
RMS error [m]	5.13	7.75	7.82	9.24
Room accuracy [%]	26.67	6.70	13.45	9.56
Response time [ms]	1 503.1	1 507.6	480.6	460.9

in Section 3.6.4), with an energy consumption of about 31 mW for the mobile node.

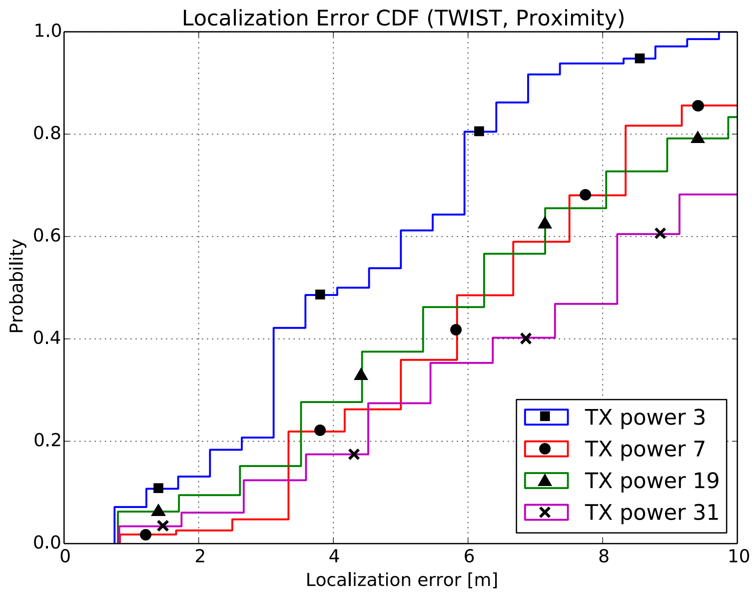


Figure 3.18: Distribution of the RMS localization error in TWIST testbed

To analyze the spatial distribution of the errors, a boxplot of the accuracy per measurement point is shown in Figure 3.19. The overall performance is for each measurement point the same, there are no obvious outliers. Noticeable, the worst minimum values are obtained in the corridor (measurement points 283, 285 and 286) and in the room where no LoS nodes are available (measurement point 240). If the results of these rooms are excluded in the room accuracy calculation, then the results are marginally better (e.g. for a tx power = 7 the room accuracy increases to 21.3% instead of to 10.3%), but even in the rooms where the nodes were available, the average error distance of almost 5 m is not enough to guarantee room accuracy: only 33.8% of all the measurement points are in the same room. The main reason for these results is that proximity requires extremely low transmission powers: even using the lowest transmission powers from the TMoteSky nodes, signals still easily penetrated the walls. Finally, the boxplot of measurement point 220 is also remarkable. The most logical explanation for this result is that only a few fixed nodes received the beacons of the mobile node. As a result, the calculator does not have much data to process. This makes the result very stable, but not necessarily more accurate.

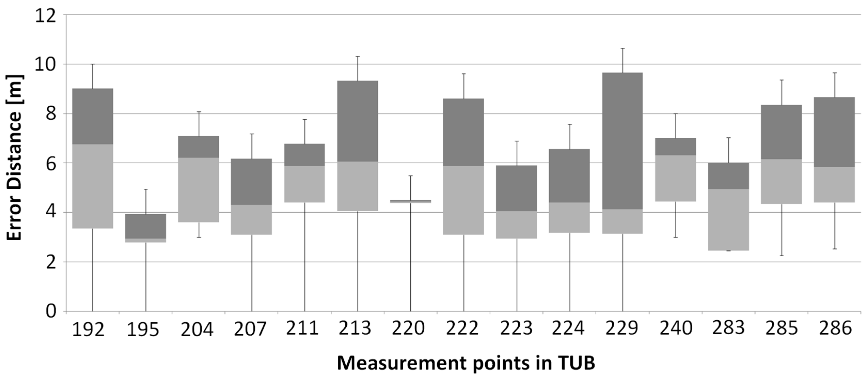


Figure 3.19: Proximity & weighted RSSI solution - boxplots representing the spatial distribution of accuracy error of each evaluation point in the TWIST environment of TUB

3.6.4 Conclusions in TWIST

An overview of the performance of the different localization solutions is given in Table 3.6.

In terms of accuracy, the best performing solutions are the fingerprinting localization techniques. Since the TWIST building represents typical office buildings with concrete and/or brick walls, different rooms are very diverse in terms of their wireless characteristics. Model based localization solutions (such as RSSI based solutions) suffer from degraded performance due to unexpected obstacles. In contrast, localization solutions that exploit this diversity, such as fingerprinting based approaches, obtain the highest accuracy.

When also considering other metrics, these conclusions need to be nuanced. In terms of response time, fingerprinting performs worst, due to the need to collect a minimal number of beacons. Since the beacon interval is not always configurable on already deployed access points, it is not always possible to decrease the response time when using existing off-the-shelf access points. In contrast, the ToA solutions can give location estimates in only halve of the response time (about 15 seconds versus 35 seconds), and as shown earlier in Figure 3.13, the response time of the RSSI based solutions can theoretically be reduced to about 200 msec.

In terms of energy consumption, the devices used in the fingerprinting solution consume most energy, which means that battery-powered solutions will have a low network lifetime when using IEEE 802.11 based fingerprinting. The energy consumption of the IEEE 802.15.4 devices is significantly lower. However, due to the large number of measurements required, the ToA still consumes twice the energy of the hybrid solution. It is clear that the energy consumption could be further optimized, albeit at the cost of longer response times.

Finally, the fingerprinting approach, although the most accurate, has one other disadvantage which is not taken into account by considering only the shown metrics. More specifically, the need for an off-line training phase and the need for retraining if the environmental conditions change can significantly impact the accuracy over time in realistic conditions. This clearly shows the need for an objective comparison method that takes into account multiple evaluation criteria when comparing localization solutions.

3.7 Results in w-iLab.t I

3.7.1 RSSI and ToA with Particle Filter

The results for the particle filter based localization using RSSI and ToA range measurements are presented in Table 3.7. Measurements are collected and processed in the same way as in the TWIST testbed.

Figure 3.20 shows how the number of measurement affects the accuracy. As in TWIST, ToA benefits from more measurements, while RSSI based ranging does not. A major difference to TWIST, also shown by the CDFs to the left in Figure 3.21, is that both types of range measurements have much larger errors. A reason for this can be that more testbed nodes far away from the measurement points are reachable, and that the far traveling signals are subject to multi-path effects to a greater extent, resulting in unpredictable attenuation that is not captured by the free-space model. The response time is also much higher than in TWIST. This is also because more testbed nodes are used for each measurement. The response time can be reduced if no measurements are collected after a certain number of testbed nodes have been used.

The right graph in Figure 3.21 shows the CDF for the localization error. The median error is about the same as that for the range measurements

Table 3.6: TWIST (Berlin) Summarized benchmarking results

Algorithm	Average Error [m]	Room accuracy [%]	Response time [ms]	Energy efficiency [mW] Mobile	Fixed
Particle filter solution					
Spray RSSI	4.35	45.00	14 285	$\sim$ 105	$\sim$ 105
Spray ToA	5.56	30.00	14 282	$\sim$ 105	$\sim$ 105
Fingerprinting solution					
KL Distance	2.7	50.0	$\sim$ 35 000	$\sim$ 7000	$\sim$ 500
ED Distance	2.2	80.0	$\sim$ 35 000	$\sim$ 7000	$\sim$ 500
PH Distance	2.0	85.0	$\sim$ 35 000	$\sim$ 7000	$\sim$ 500
Hybrid solution					
TX Power = 3	4.6	26.7	1 503.1	$\sim$ 30.9	$\sim$ 47.4
TX Power = 7	7.1	6.7	1 507.6	$\sim$ 35.1	$\sim$ 47.4
TX Power = 19	7.9	13.4	480.6	$\sim$ 47.1	$\sim$ 47.4
TX Power = 31	8.7	9.5	460.9	$\sim$ 57.6	$\sim$ 47.4

Table 3.7: Accuracy of the particle filter in the w-iLab.t I testbed

	RSSI	**ToA**
Average error [m]	7.79	7.16
Min. error [m]	3.59	1.51
Max. error [m]	14.04	14.31
Median error [m]	7.09	6.09
RMS error [m]	8.43	7.92
Room accuracy [%]	30.00	20.00
Response time [ms]	55,448	55,444

in the left graph (8 and 7 m, respectively), but has lower errors above the median.

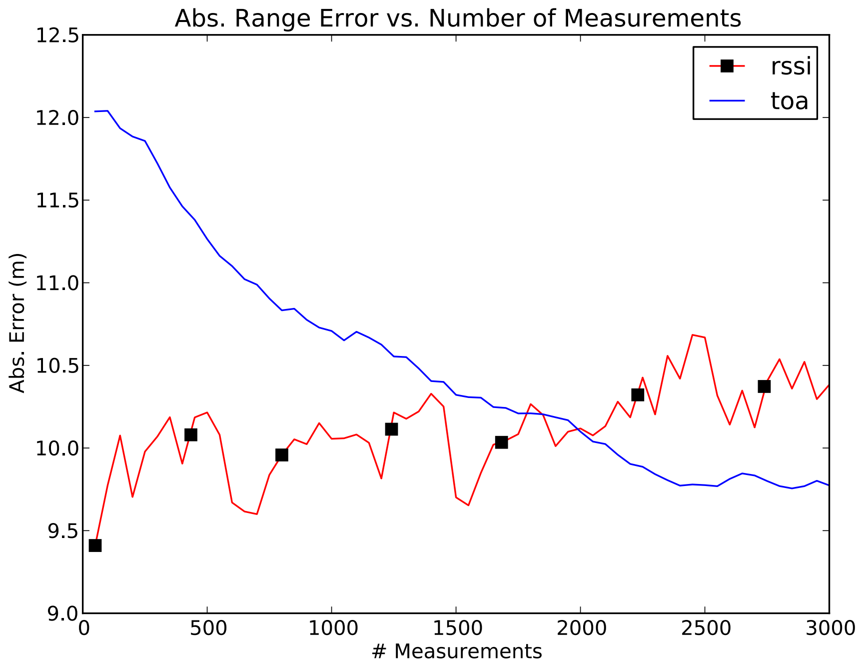


Figure 3.20: The absolute range error for ToA decreases with the number of measurements until approximately 2500 measurements. The RSSI error fluctuates about the same value, and is not improved by additional measurements.

The power consumption for the mobile node is the same as in the TWIST experiments, i.e. 105 mW. Although less testbed nodes are used in this experiment, we consider the total infrastructure power consumption to be of the same magnitude as in TWIST, i.e. 7 W.

3.7.2 Fingerprinting

This section evaluates the fingerprinting localization approach described in Section 3.3.2 in the w-iLab.t I testbed. The location of the IEEE 802.11 Wi-Fi access points used to create fingerprints can be found in Figure 3.22, marked with red dots.

The accuracy results are shown in Table 3.8. In general, the accuracy is lower than in the TWIST environment. The CDF of localization error is shown in Figure 3.23. The decrease in accuracy can be explained by the fact that w-iLab.t I is an office environment that uses plywood walls,

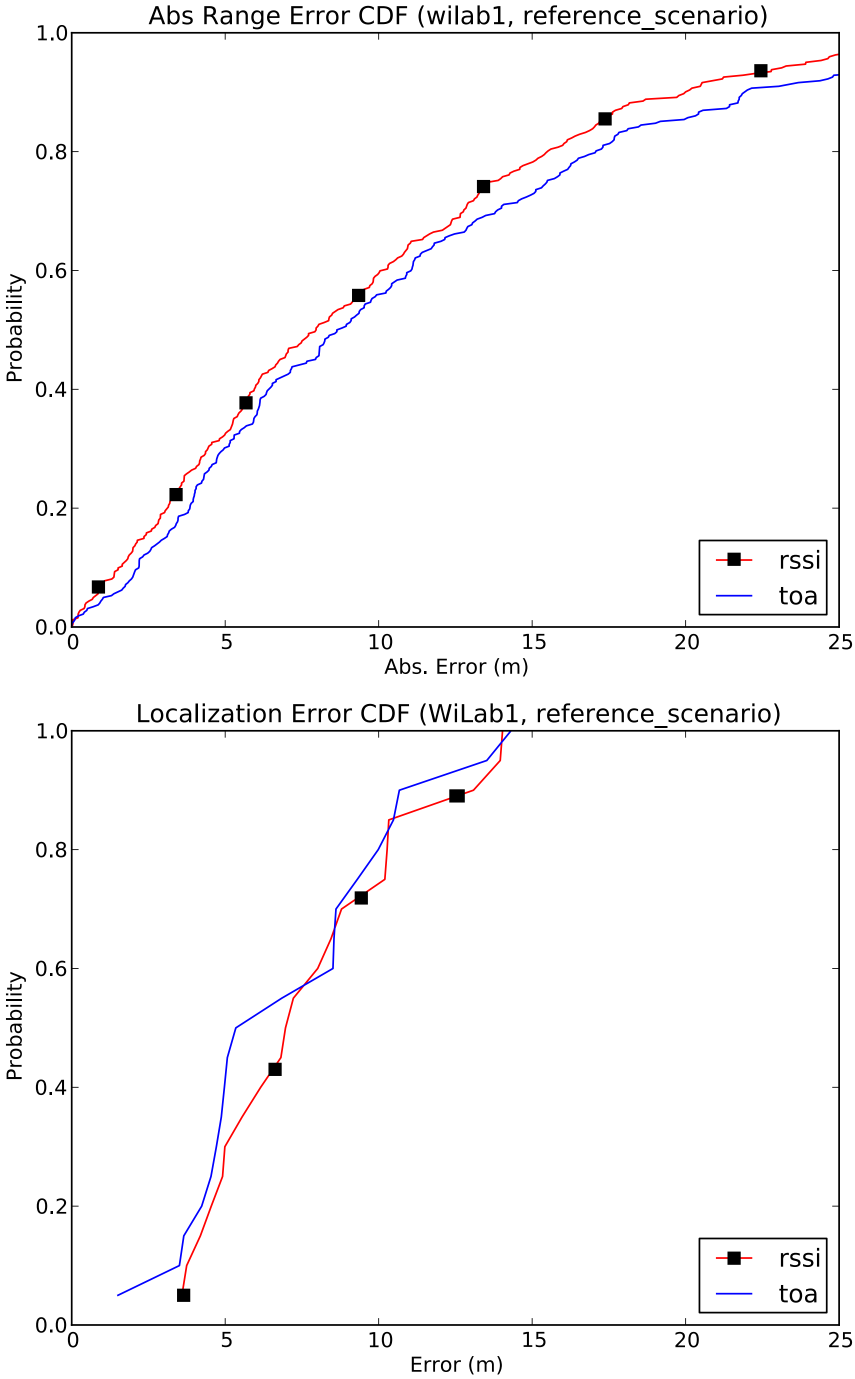


Figure 3.21: CDFs for the absolute range error (top) and the localization error (bottom).

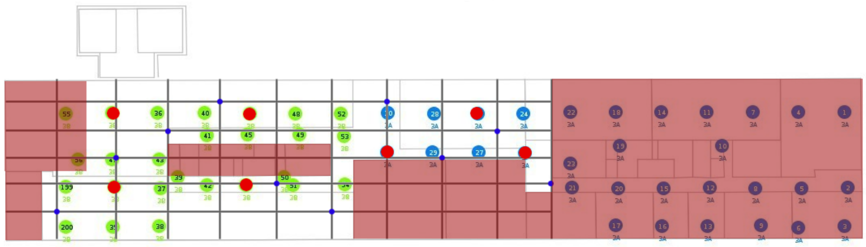


Figure 3.22: w-iLab.t I: IEEE 802.11 access points

which attenuate the signals less than the concrete walls from TWIST. As a result, the evaluated locations have less diversity in terms of the received signal strengths, and are thus more difficult to uniquely characterize in a fingerprint. This effect will have an even greater influence in the results of open room environment of w-iLab.t II (Section 3.8). The point and room level accuracy vary with the performance delay as presented in Figure 3.24 and 3.25.

Table 3.8: Statistical information about the fingerprinting algorithms' performance in w-iLab.t I

	KL	ED	PH
Average error [m]	6.15	2.37	2.75
Min. error [m]	1.12	0.00	0.00
Max. error [m]	15.86	5.50	11.0
Median error [m]	4.37	2.75	2.75
RMS error [m]	7.25	3.34	3.76
Room accuracy [%]	50.0	80.0	85.0
Response time [s]	24.98	24.16	24.36

3.7.3 Proximity & weighted RSSI

The location accuracy of the weighted RSSI based localization solution in w-iLab.t I is shown in Table 3.9. The CDF of the localization error can be found in Figure 3.26. Because the plywood walls don't attenuate the signals significantly, locations need to be determined based on weighted RSSI values (rather than proximity) even when using low transmission powers. As a result, in contrast to the experiments in the TWIST testbed where the localization accuracy depends strongly on the transmission power, the results in w-iLab.t I are less dependent on the transmission power. The spatial spread of the accuracy is shown in Figure 3.27 using a boxplot.

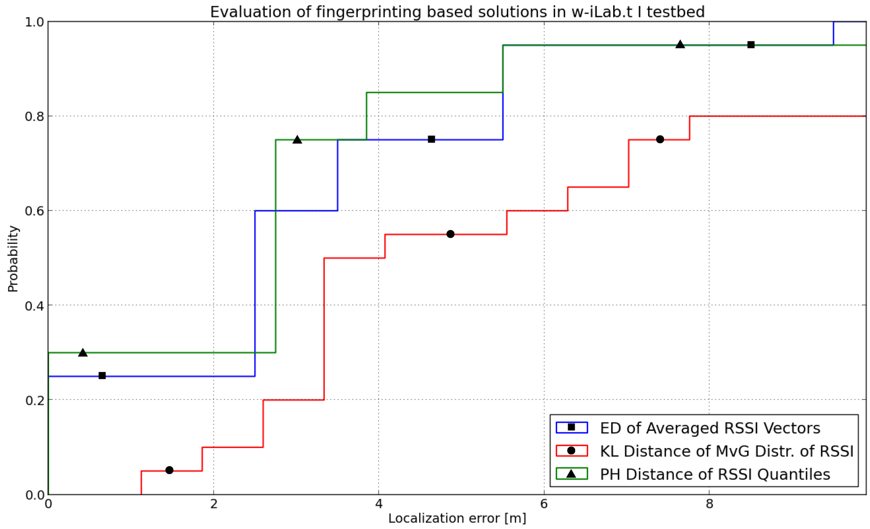


Figure 3.23: CDF of the localization error of fingerprinting based solutions in w-iLab.t I testbed

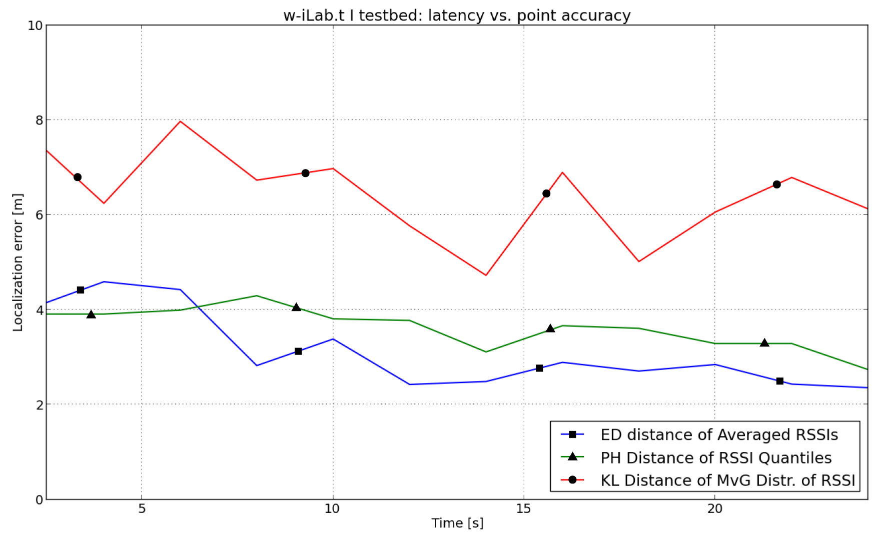


Figure 3.24: Fingerprint collection delay versus point accuracy

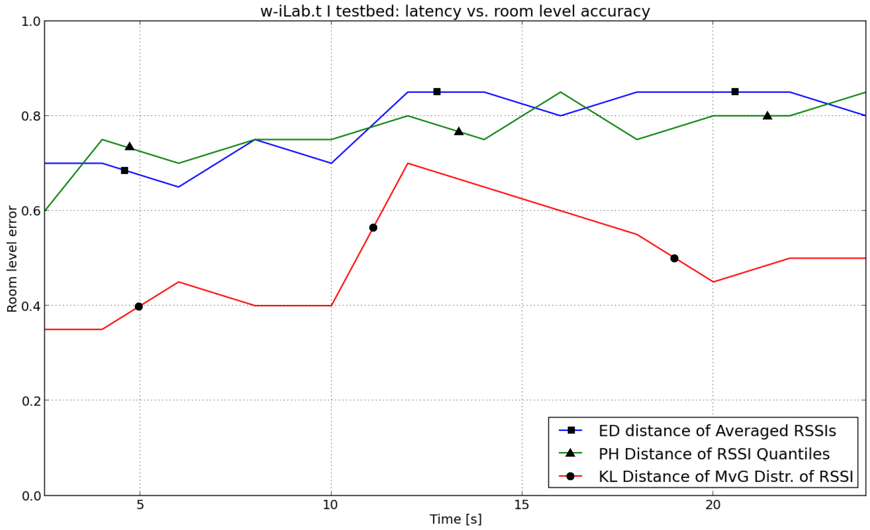


Figure 3.25: Fingerprint collection delay versus room level accuracy

The measured points more in the center of the testbed have a higher accuracy then those at the edges, because the edge evaluation points are outside the grid of used anchor points. For example, in Figure 3.28 a clear bias in the estimated locations can be observed caused by the fact that all anchor nodes are located at the same side of the evaluation point. This highlights the importance of using anchor nodes outside the area that is evaluated, which is a requirement that is not found for the fingerprinting solutions. Finally, Figure 3.29 shows the room accuracy. It is interesting to note that it is not possible to predict the room accuracy based only on the point accuracy, because the room accuracy depends strongly on random factors such as the direction of the inaccuracies.

Table 3.9: Statistical information about the hybrid algorithms' performance in w-iLab.t I

	Tx3	Tx7	Tx19	Tx31
Average error [m]	7.64	8.86	7.47	8.21
Min. error [m]	0.07	0.04	0.17	0.04
Max. error [m]	48.77	65.98	45.23	45.23
Median error [m]	5.87	7.63	6.18	7.21
RMS error [m]	9.35	10.15	8.83	9.44
Room accuracy [%]	18.62	16.27	12.60	9.46
Response time [ms]	2100.74	113.15	107.99	110.17

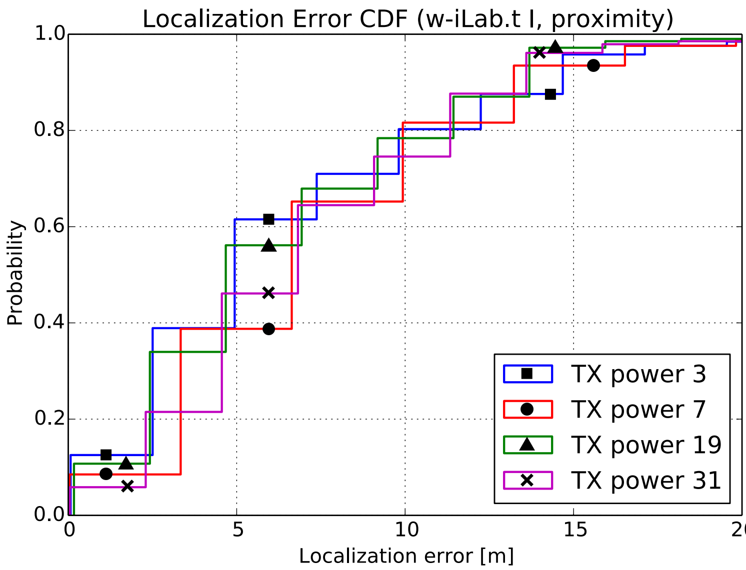


Figure 3.26: Distribution of the RMS localization error in w-iLab.t I testbed

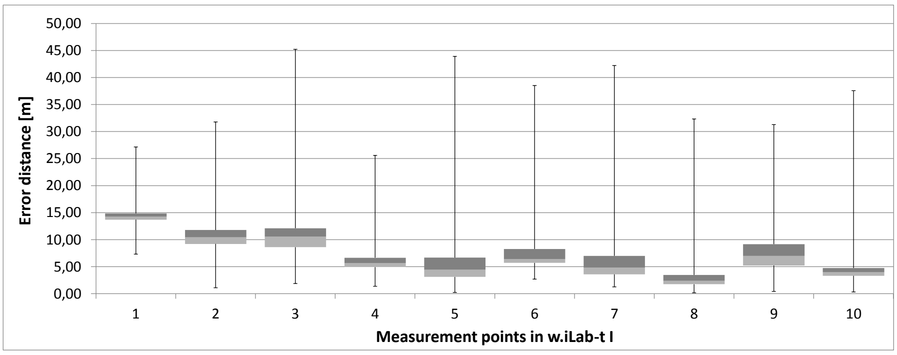


Figure 3.27: Proximity & weighted RSSI solution - spatial distribution of accuracy error in w-iLab.t I, including maximum error, minimum error, quartile 1, quartile 2 and median error

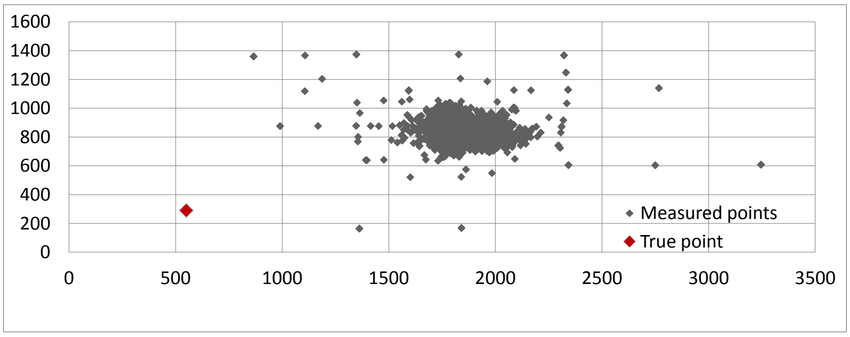


Figure 3.28: Biased spread of the location estimates resulting from evaluating measurement locations outside the grid of anchor nodes.

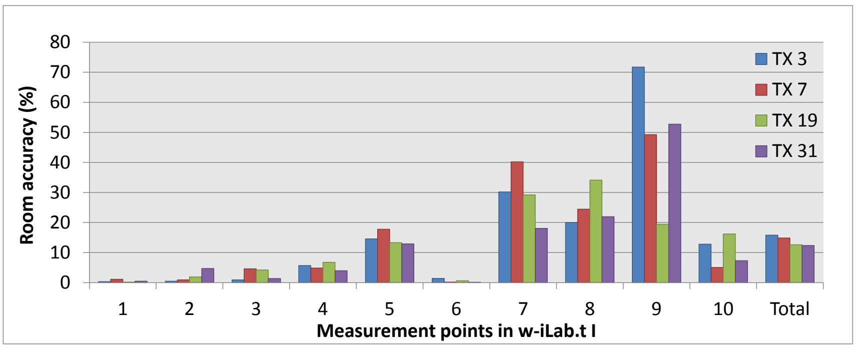


Figure 3.29: The results of the room accuracy in each measurement point in the w-iLab.t testbed.

3.7.4 Conclusions from the w-iLab.t I experiments

An overview of the performance of the different localization solutions is given in Table 3.10. As a representative of a typical building with plywood walls, signals are less attenuated than in the TWIST testbed, resulting in less unique wireless features per room. As a result, fingerprinting solutions perform worse than in the TWIST environment.

Also the ToA and RSSI based solutions have significantly degraded performance. This can be explained by the fact that, due to testbed limitations, anchor nodes are not installed on the corner points, meaning that several evaluation points are outside the grid of anchor nodes. In addition, although the walls are made from plywood that have a very small attenuation factor, signal propagation still behaves very unpredictable due to the presence of large metal cupboards and metal ceilings. This demonstrates that the performance of localization solutions in typical environments is influenced by many factors besides the building construction materials, and highlights the fact that localization performances measured in a empty building should not be considered representative for the performance of said solutions when the buildings are actively used.

3.8 Results in w-iLab.t II

3.8.1 RSSI and ToA with Particle Filter

The results for the particle filter based localization using RSSI and ToA range measurements are presented in Table 3.11.

In this testbed, the measurements are collected in a slightly different way than in TWIST and w-iLab.t I. A single channel is used due to a limitation of the testbed nodes. Moreover, instead of collecting multiple measurements from a specific testbed node before switching to the next testbed node, a single message is exchanged with each infrastructure node and when all nodes have been tried, the process starts over again with the first node. As a result, measurements are collected from more testbed nodes, but each having fewer measurements. For this reason a maximum of approximately 900 measurements are collected from a single node in this testbed. Figure 3.30 shows that, at least for within this range of collected measurements, the accuracy is not affected for any of the ranging methods. Although few measurements, are collected from each node, we observe a high response time due to the fact that many testbed nodes are used at each measurement point. As in the w-iLab.t I testbed, the response time can be reduced by limiting the amount of testbed nodes used at each measurement point.

Figure 3.31 shows the CDFs for the range measurements and the localization estimations. Although the ToA range measurements (left graph)

Table 3.10: w-iLab.t I (Ghent) Summarized benchmarking results

Algorithm	Average Error [m]	Room accuracy [%]	Response time [ms]	Energy efficiency [mW] Mobile	Energy efficiency [mW] Fixed
Particle filter solution					
Spray RSSI	7.79	30.00	55 448	~105	~105
Spray ToA	7.16	20.00	55 444	~105	~105
Fingerprinting solution					
KL Distance	6.15	50.00	~24000	~7000	~500
ED Distance	2.37	80.00	~24000	~7000	~500
PH Distance	2.75	85.00	~24000	~7000	~500
Hybrid solution					
TX Power = 3	7.64	18.62	2100.74	~30.9	~47.4
TX Power = 7	8.86	16.27	113.15	~35.1	~47.4
TX Power = 19	7.47	12.60	107.99	~47.1	~47.4
TX Power = 31	8.21	9.46	110.17	~57.6	~47.4

Table 3.11: Statistical information about the particle filter algorithms' performance in w-iLab.t II

	RSSI	**ToA**
Average error [m]	6.41	6.66
Min. error [m]	0.90	0.99
Max. error [m]	20.22	27.06
Median error [m]	5.68	5.50
RMS error [m]	8.05	8.59
Response time [ms]	59633	59620

are less accurate than that of the RSSI based method, the final localization estimations (right graph) for the two methods are more or less equal.

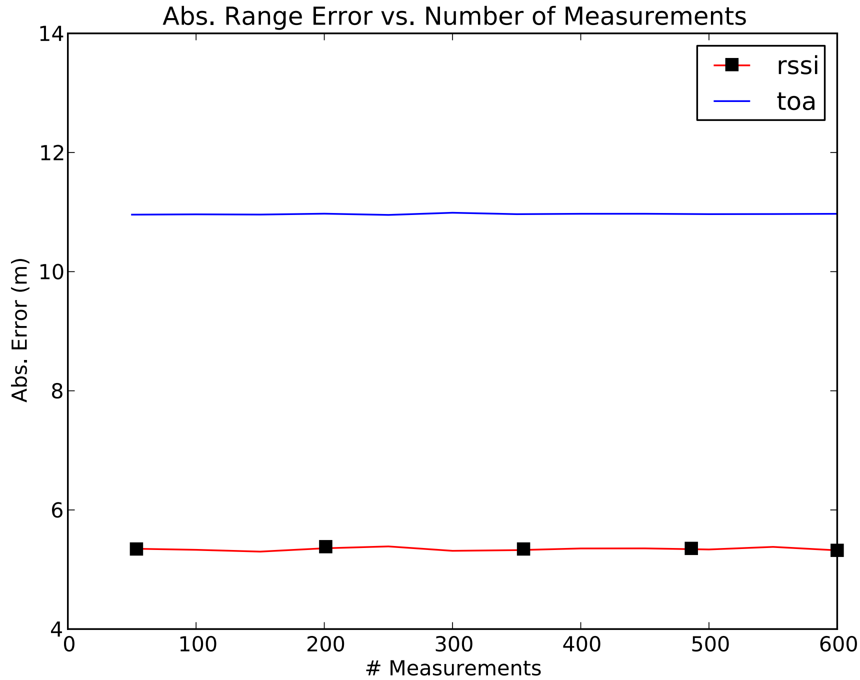


Figure 3.30: The absolute range error is not affected for neither ToA nor RSSI measurements.

The power consumptions stated for TWIST and w-iLab.t I are also applicable here.

3.8.2 Fingerprinting

Table 3.12 contains the accuracy statistics of the fingerprinting localization solutions described in Section 3.3.2. No room accuracy is reported because the testbed consists of a single large open space. The accuracy is significantly lower than the accuracy obtained in the other testbeds (Figure 3.32). This degradation is mainly caused by two physical characteristics of the environment: (i) no separate rooms are present which makes it difficult to create unique fingerprints for each location and (ii) due to the metal walls, random reflection result in signal strengths that vary strongly from packet to packet.

Figure 3.33 shows the influence of collecting additional data before creating fingerprints. It demonstrates the importance of using robust finger-

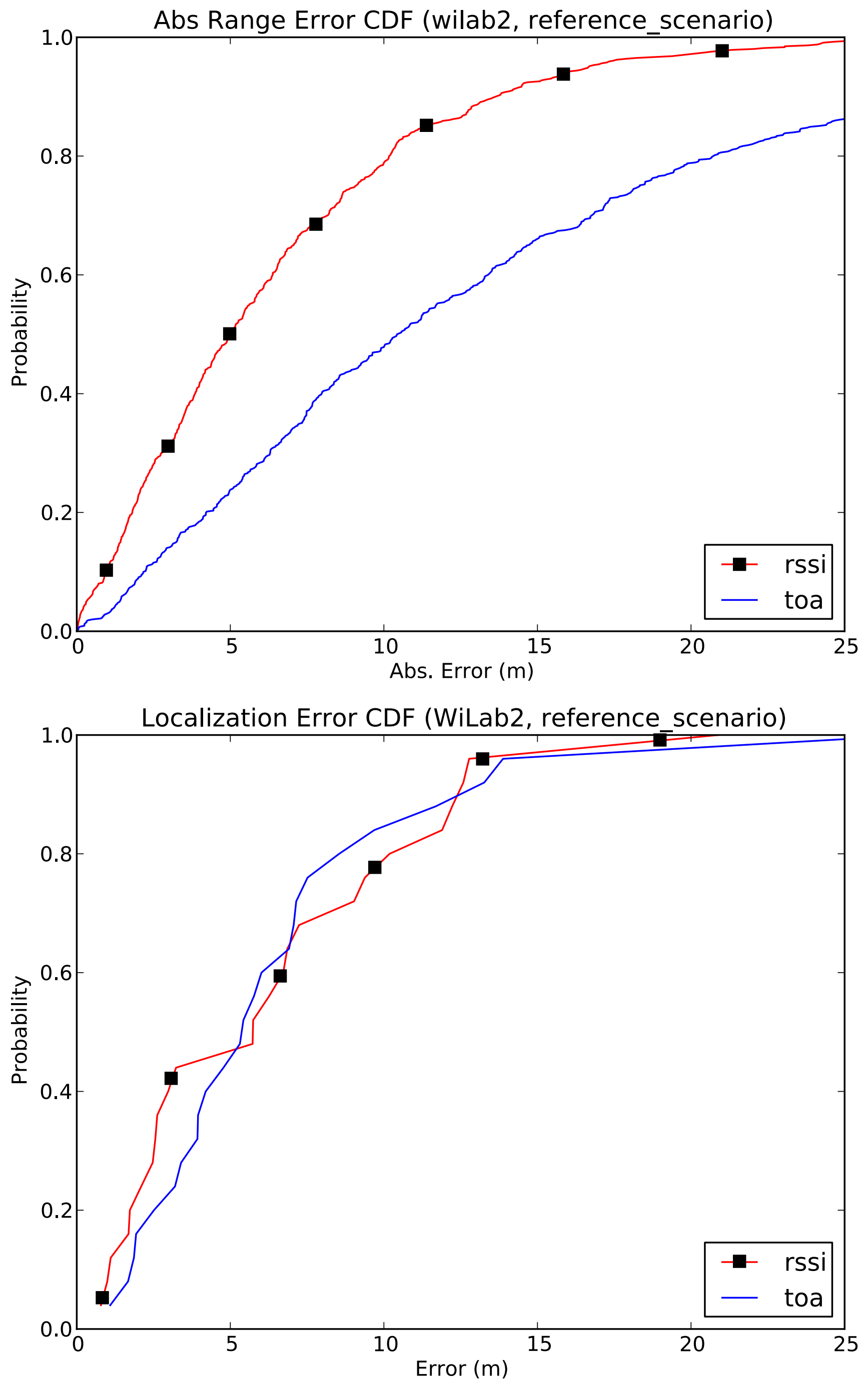


Figure 3.31: CDFs for the absolute range error (top) and the localization error (bottom).

Table 3.12: Statistical information about the performance of fingerprinting algorithms in w-iLab.t II testbed

	KL	ED	PH
Average error [m]	24.76	19.08	8.13
Min. error [m]	3.00	3.00	0.00
Max. error [m]	47.43	39.00	15.10
Median error [m]	21.0	18.97	6.70
RMS error [m]	28.09	20.76	8.97
Response time [s]	24.78	24.37	24.12

printing creation methods (e.g. PH distance of RSSI quantiles) and demonstrates that these robust fingerprinting creation methods can be used to generate more accurate results, on the condition that more data is collected (at the cost of higher response delays).

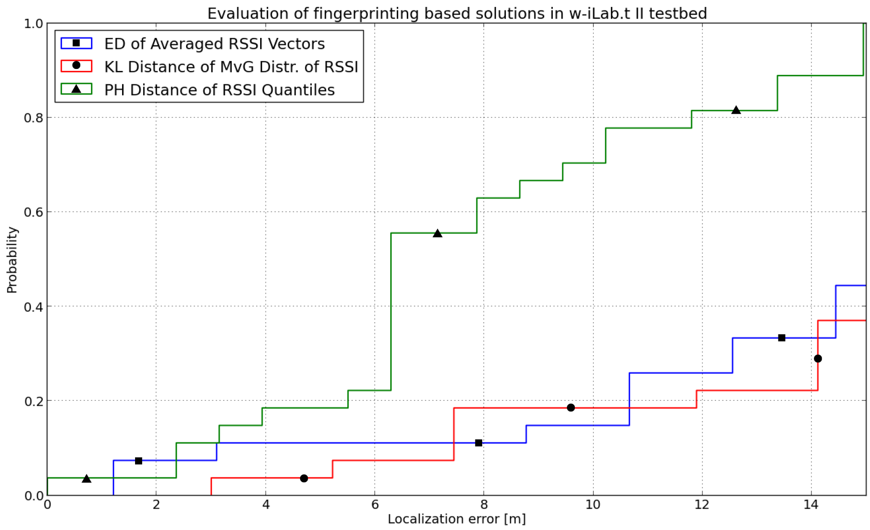


Figure 3.32: CDF of the localization error of fingerprinting based solutions in w-iLab.t II testbed

3.8.3 Proximity & weighted RSSI

The location accuracy of the weighted RSSI based localization solution in w-iLab.t II is shown in Table 3.13. The CDF of the localization error can be found in Figure 3.34. The average accuracy is significantly lower than in the previous environments, mainly due to the many reflections in the

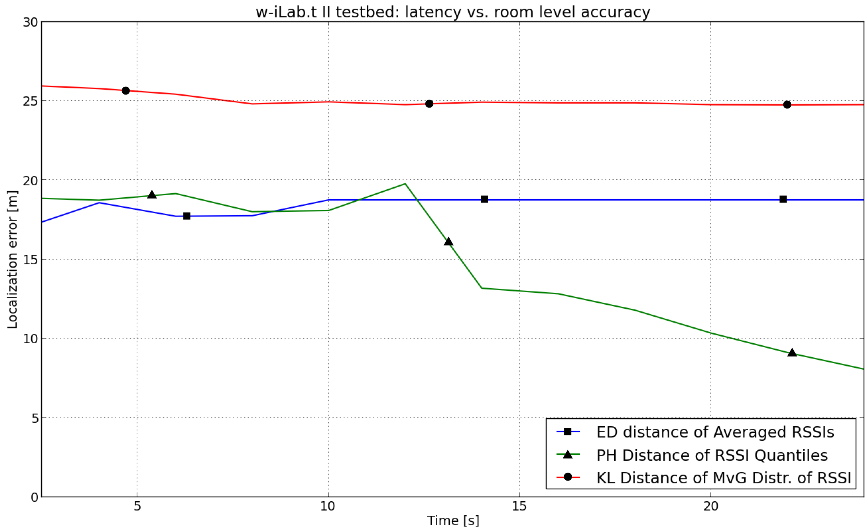


Figure 3.33: Fingerprint collection delay versus point accuracy

environment thereby causing self-interference. The spatial spread of the accuracy is shown in Figure 3.35 using a boxplot.

Table 3.13: Accuracy of the hybrid localization algorithm in the w-iLab.t II testbed (TX power 31)

Average error [m]	17.16
Min. error [m]	1.57
Max. error [m]	52.15
Median error [m]	16.24
RMS error [m]	19.73
Response time [ms]	15.70

3.8.4 Conclusions from the w-iLab.t II experiments

The environment from w-iLab.t II exhibits characteristics which are typical for many large-size industrial indoor environments, namely open spaces surrounded by metal obstacles, walls and ceilings. The results clearly indicate that all tested types of RF-based localization solutions degrade significantly in these environments. All the signals have a lot of reflections with the metal construction, causing a lot of multipath effects. This indicates that accurate indoor localization in industrial open environments is a difficult task.

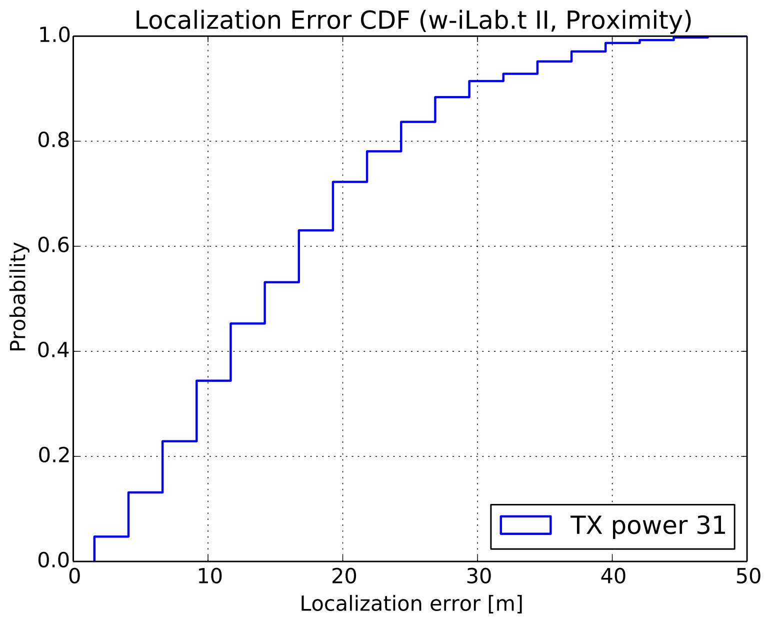


Figure 3.34: Distribution of the RMS localization error in w-iLab.t II testbed

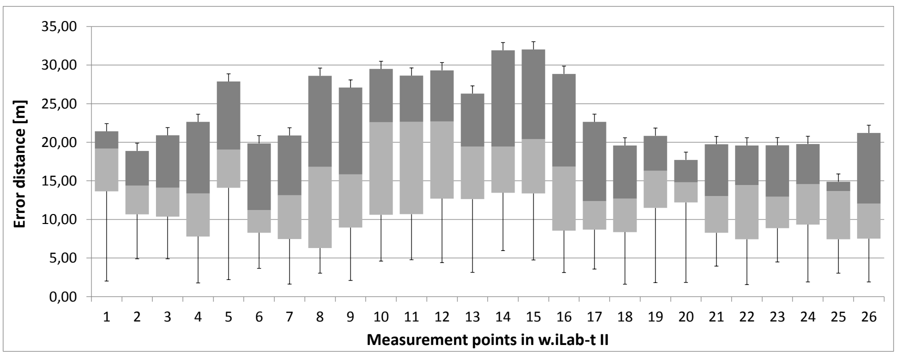


Figure 3.35: Proximity & weighted RSSI solution - spatial distribution of accuracy error in w-iLab.t II, including maximum error, minimum error, quartile 1, quartile 2 and median error

In this industrial environment, the ToA and RSSI based ranging solutions using the particle filter contains the best results. An average error distance of 6m - 7m instead of 11m for the Fingerprinting and 17m for the hybrid technique.

Table 3.14: w-iLab.t II (Zwijnaarde) Summarized benchmarking results

Algorithm	Average Error [m]	Room accuracy [%]	Response time [ms]	Energy efficiency [mW] Mobile	Energy efficiency [mW] Fixed
Particle filter solution					
Spray RSSI	6.41	-	59 633	∼105	∼105
Spray ToA	6.66	-	59 620	∼105	∼105
Fingerprinting solution					
KL Distance	24.76	-	∼24 000	∼7 000	∼500
ED Distance	19.08	-	∼24 000	∼7 000	∼500
PH Distance	8.13	-	∼24 000	∼7 000	∼500
Hybrid solution					
TX Power = 31	17.16	-	15.7	∼57.6	∼47.4

3.9 Discussion

Over the course of the performed experiments, several lessons were learned. First of all, the experiments clearly show the importance of choosing representative measurement locations. Several locations have consistent lower accuracy results. For example the hallways (which are narrow and as such have very low room accuracy). As such, it is clear that the localization points should include a representative mix of 'easy to locate positions' and more challenging ones. Ideally, a fine-grained grid-like approach should be used in which the positioning accuracy is evaluated every X meter. All experiments in this chapter use the same evaluation points.

Another lesson learned is that the location accuracy differs strongly between different testbeds. Environment specifics (such as metal ceilings) strongly influence propagation behavior. The highest accuracy was obtained in more 'traditional' brick-wall office scenarios, such as represented by the TWIST testbed. The w-iLab.t I testbed, which has plywood walls and metal ceilings, has a lower accuracy. Finally, w-iLab.t II consists of a fully shielded environment, in which the walls and ceilings are from metal, and contains a number of metal obstructions. Performing localization in this testbed, i.e.: a confined and strongly reflecting environment, proves to be very challenging.

Thirdly, the experiments and benchmarking results that are executed illustrate the need for evaluating the full subset of the metrics. Although the accuracy of the fingerprinting solutions in office environments is shown to be very good, these solutions require significantly more time to collect beacons for fingerprinting, which strongly influences the response delay. Similarly, the Wi-Fi based fingerprinting localization solutions perform very well, but have higher energy requirements than the solutions using sensor nodes.

Moreover, the experiments indicate that the performance of localization solutions strongly depends on several algorithmic and deployment aspects, such as the used technology, the ranging approach, the location estimation approach, post- and preprocessing, anchor positions, etc. Making even minor changes to one of these aspects can have a profound influence on several performance metrics. It was also shown that the internal configuration of the algorithms, such as preprocessing the data (such as removing the 10% highest and lowest outlier values) or setting the minimum number of beacons that is collected for location estimation can significantly influence the performance.

As such, to allow objective comparison of localization solutions, it is clear that independent evaluation procedures should be defined by an impartial third party and that such evaluation procedures should include at least the following aspects: (i) definition of a wide set of evaluation metrics, (ii) clear definition of the evaluation environments in which the results are valid (iii) an objective method for generating a representative set of evaluation points.

3.10 Conclusion

Although many indoor localization solutions exist, this chapter pointed out that the scientific evaluation methods for RF based solutions are currently limited in scope. Localization solutions are evaluated mainly based on location accuracy and are evaluated in a single testbed environment. As a result, it is not clear to which level the results from existing scientific literature can be compared to each other.

To evaluate how these different conditions can influence the localization performance, three localization solutions were selected that represent typical approaches for indoor localization, including multiple technologies (IEEE 802.11 and IEEE 802.15.4), multiple localization approaches (fingerprinting, time-of-arrival and RSSI-based) and multiple processing methods. To allow objective comparisons, the same evaluation methodology was used to evaluate the performance of these localization solutions in three different environments: an office environment with brick walls, an office environment with plywood walls and an open environment with metal walls and metal obstacles.

The main conclusion of these experiments were the following.

- Several inherent trade-offs between different metrics have been identified, which are typically ignored when reporting only on the accuracy of the solutions. More specifically, the results show a very clear trade-off between the collected number of measurements (which are directly translated into energy consumption and response delay) and the point-level accuracy.
- The accuracy of localization solutions depends strongly on the characteristics of the environment. Due to the presence of concrete and/or brick walls in the office testbeds, different rooms are very diverse in terms of their wireless characteristics. Localization solutions that exploit this diversity, such as fingerprinting based approaches, obtain the highest accuracy. In contrast, in more industrial-like open environments time-of-arrival solutions performed better. These results show that future scientific literature describing performance results of localization solutions should include detailed descriptions of the used evaluation environment(s), including information such as propagation characteristics, typical room sizes and a description of the materials used in walls and ceilings.
- We have shown that the choice of evaluation points strongly influence the reported accuracy. As such, papers that use self-selected evaluation points can significantly influence their reported accuracy by artificially selecting those evaluation points that outperform other locations.
- Due to testbed constraints, one evaluation environment contained evaluation points outside the grid of anchor points. It was shown that this set-up had a negative influence on some of the evaluated solutions (mainly the RSSI-based and time-of-arrival solutions) but not on the fingerprinting solution. As such, when considering which is the best localization solution for an industrial deployment, building layout constraints should also be included.

- The accuracy can decrease significantly when evaluating an environment for which the localization solution is not specifically tweaked. For example, all tested solutions suffered from degraded accuracy in the open industrial-like environment, up to a factor 10 lower. Since most existing solutions have been optimized for office environments, these results hint that many existing localization solutions might not be ready for use in industrial environments or other challenging environments, such as underground mines.

The above findings reveal several weaknesses in the evaluation methods used in the majority of existing scientific nature of indoor localization solutions. As such, there is a clear need for a standardized evaluation methodology to objectively compare different localization solutions in multiple conditions.

References

[1] N. Bulusu, J. Heidemann, and D. Estrin. *GPS-less low-cost outdoor localization for very small devices*. Personal Communications, IEEE, 7(5):28–34, 2000.

[2] J. Hightower and G. Borriello. *Location systems for ubiquitous computing*. Computer, (8):57–66, 2001.

[3] K. Pahlavan, X. Li, and J.-P. Mäkelä. *Indoor geolocation science and technology*. Communications Magazine, IEEE, 40(2):112–118, 2002.

[4] H. Liu, H. Darabi, P. Banerjee, and J. Liu. *Survey of wireless indoor positioning techniques and systems*. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37(6):1067–1080, 2007.

[5] T. Van Haute, E. De Poorter, F. Lemic, and W. Niklas. *D2.2: Report on experiments without interference*. Public deliverable, The EVARILOS Project (grant agreement no 317989.), 2013.

[6] *The EVARILOS project*. http://www.evarilos.eu. Accessed: 2016-05-20.

[7] T. Van Haute, E. De Poorter, J. Rossey, I. Moerman, V. Handziski, A. Behboodi, F. Lemic, A. Wolisz, N. Wiström, T. Voigt, et al. *The evarilos benchmarking handbook: Evaluation of rf-based indoor localization solutions*. In 2e International Workshop on Measurement-based Experimental Research, Methodology and Tools (MERMAT 2013), pages 1–6, 2013.

[8] T. Van Haute, E. De Poorter, F. Lemic, and W. Niklas. *D2.1: EVARILOS Benchmarking Handbook*. Public deliverable, The EVARILOS Project (grant agreement no 317989.), 2013.

[9] F. Lemic, N. Wirström, and V. Handziski. *D2.3: Report on experiments with controlled interference*. Public deliverable, The EVARILOS Project (grant agreement no 317989.), 2013.

[10] *ISO Standard - ISO/IEC CD 18305 - Information technology – Real time locating systems – Test and evaluation of localization and tracking systems*. http://www.iso.org/iso/home/store/catalogue_htc/catalogue_detail.htm?csnumber=62090. Accessed: 2016-05-20.

[11] *EvAAL - Evaluating AAL Systems through Competitive Benchmarking*. http://evaal.aaloa.org/index.php?option=com_content&view=article&id=187:technical-annexes-localization2013&catid=15&Itemid=261. Accessed: 2016-05-20.

[12] D. Yao, C. Yu, A. K. Dey, C. Koehler, G. Min, L. T. Yang, and H. Jin. *Energy efficient indoor tracking on smartphones*. Future Generation Computer Systems, 39:44–54, 2014.

[13] V. Otsason, A. Varshavsky, A. LaMarca, and E. De Lara. *Accurate GSM indoor localization*. In UbiComp 2005: Ubiquitous Computing, pages 141–158. Springer, 2005.

[14] A. M. Ladd, K. E. Bekris, A. P. Rudys, D. S. Wallach, and L. E. Kavraki. *On the feasibility of using wireless ethernet for indoor localization*. IEEE Transactions on Robotics and Automation, 20(3):555–559, 2004.

[15] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan. *Indoor localization without the pain*. In Proceedings of the sixteenth annual international conference on Mobile computing and networking, pages 173–184. ACM, 2010.

[16] E. Martin, O. Vinyals, G. Friedland, and R. Bajcsy. *Precise indoor localization using smart phones*. In Proceedings of the international conference on Multimedia, pages 787–790. ACM, 2010.

[17] S. Woo, S. Jeong, E. Mok, L. Xia, C. Choi, M. Pyeon, and J. Heo. *Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR*. Automation in Construction, 20(1):3–13, 2011.

[18] M. Stella, M. Russo, and D. Begušić. *Fingerprinting based localization in heterogeneous wireless networks*. Expert systems with applications, 41(15):6738–6747, 2014.

[19] C. Steiner and A. Wittneben. *Efficient training phase for ultrawideband-based location fingerprinting systems*. Signal Processing, IEEE Transactions on, 59(12):6021–6032, 2011.

[20] P. Pettinato, N. Wirström, J. Eriksson, and T. Voigt. *Multi-channel two-way time of flight sensor network ranging*. In Wireless Sensor Networks, pages 163–178. Springer, 2012.

[21] J. Elson and K. Römer. *Wireless sensor networks: A new regime for time synchronization*. ACM SIGCOMM Computer Communication Review, 33(1):149–154, 2003.

[22] E. B. Mazomenos, D. De Jager, J. S. Reeve, and N. M. White. *A two-way time of flight ranging scheme for wireless sensor networks*. In Wireless Sensor Networks, pages 163–178. Springer, 2011.

[23] N. Wirstrom, P. Misra, and T. Voigt. *Spray: a multi-modal localization system for stationary sensor network deployment*. In Wireless On-demand Network Systems and Services (WONS), 2014 11th Annual Conference on, pages 25–32. IEEE, 2014.

[24] F. Lemic. *Benchmarking of Quantile based Indoor Fingerprinting Algorithm*. Telecommunication Networks Group, Technische Universität Berlin, Tech. Rep. TKN-14-001, 2014.

[25] D. Milioris, G. Tzagkarakis, A. Papakonstantinou, M. Papadopouli, and P. Tsakalides. *Low-dimensional signal-strength fingerprint-based positioning in wireless LANs*. Ad hoc networks, 12:100–114, 2014.

[26] J. M. Chambers, D. A. James, D. Lambert, and S. V. Wiel. *Monitoring networked applications with incremental quantile estimation*. Statistical Science, pages 463–475, 2006.

[27] T. Van Haute, J. Rossey, P. Becue, E. De Poorter, I. Moerman, and P. Demeester. *A hybrid indoor localization solution using a generic architectural framework for sparse distributed wireless sensor networks*. In Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on, pages 1009–1015. IEEE, 2014.

[28] C.-N. Huang and C.-T. Chan. *ZigBee-based indoor location system by k-nearest neighbor algorithm with weighted RSSI*. Procedia Computer Science, 5:58–65, 2011.

4

Performance Analysis of Multiple Indoor Positioning Systems in a Healthcare Environment

This chapter is comparable with the previous one, we perform a comprehensive performance analysis of three different types of indoor localization solutions (a Wi-Fi based fingerprinting, a Zigbee based Time of Arrival (ToA) solution and a Zigbee based multilateration algorithm) in a healthcare environment. However there are crucial differences: (i) these experiments are evaluated in a real-life environment instead of a dedicated environment for testing purposes. (ii) The previous chapter its main focus was to discover and to prove the lack of comparability, whilst this chapter its main goal was to evaluate and optimize the initial version of the benchmarking methodology and platform. Consequently, this chapter also elaborates the procedure for acquiring datasets in a real-life environment which can be reused by multiple solutions.

⋆ ⋆ ⋆

T. Van Haute, E. De Poorter, P. Crombez, F. Lemic, V. Handziski, N. Wirström, A. Wolisz, T. Voigt, I. Moerman.

Published in International Journal of Healthcare Geographics, February 2016. [IF: 2.447 , Q1]

Abstract The combination of an aging population and nursing staff shortages implies the need for more advanced systems in the healthcare industry. Many key enablers for the optimization of healthcare systems require provisioning of location awareness for patients (e.g. with dementia), nurses, doctors, assets, etc. Therefore, many Indoor Positioning Systems (IPSs) will be indispensable in healthcare systems. However, although many IPSs have been proposed in literature, most of these have been evaluated in non-representative environments such as office buildings rather than in a hospital.

To remedy this, the paper evaluates the performance of existing IPSs in an operational modern healthcare environment: the “Sint-Jozefs kliniek Izegem” hospital in Belgium. The evaluation (data-collecting & data-processing) is executed using a standardized methodology and evaluates the point accuracy, room accuracy and latency of multiple IPSs. To evaluate the solutions, the position of a stationary device was requested at 73 evaluation locations. By using the same evaluation locations for all IPSs the performance of all systems could objectively be compared.

Several trends can be identified such as the fact that Wi-Fi based fingerprinting solutions have the best accuracy result (point accuracy of $1.21\ m$ and room accuracy of $98\ \%$) however it requires calibration before use and needs $5.43\ s$ to estimate the location. On the other hand, proximity based solutions (based on sensor nodes) are significantly cheaper to install, do not require calibration and still obtain acceptable room accuracy results.

As a conclusion of this paper, Wi-Fi based solutions have the most potential for an indoor positioning service in case when accuracy is the most important metric. Applying the fingerprinting approach with an anchor installed in every two rooms is the preferred solution for a hospital environment.

4.1 Introduction

In recent years, the complexity in nursing facilities has been increasing due to societal factors such as the increase of the care unit size, the increase of specialized care and the lack of nurse staffing, which requires a more efficient use of resources [1]. In addition to these inherent factors, a further increase in complexity is due to different technologies that is being introduced for the staff (e.g. medical equipment, pagers, alert redirecting and electronic medical records) as well as for the environment (e.g. building automation for energy control and comfort functions for the patients). In future years, these complexity trends will continue due to upcoming technologies (such as location aware services and computerized decision support systems) and an ageing society [2], which translates into an increasing need for care and a decrease of the available staff.

The introduction of location awareness in healthcare environments rises a wide range of new possibilities [3]. An Indoor Positioning System (IPS) allows hospitals to locate persons or assets inside the building. Interesting hospital scenarios could become reality: advanced nurse calling systems could locate the nearest nurse, making their work more efficient [4]. Patients with dementia will experience more freedom since they should not be locked away anymore. As a final example: finding assets inside a building can be a complicated task. In many cases time matters, finding an important asset faster can save lives. These are only a few examples how an IPS can improve the internal functionalities inside a hospital.

To avoid confusion, it is important to make a distinction between a "positioning" and a "tracking" system. The latter uses history based information to estimate the location of the person or asset that needs to be tracked. This implies multiple negative consequences: (i) a start reference point is crucial when tracking is involved. If this is not calibrated carefully, tracking results will be useless. (ii) The mobile node needs to communicate continuously to keep the location updated which will drain the battery much faster. (iii) Finally, this causes conflicts in terms of privacy. When doctors or nurses are being tracked, their entire location history is available. Due to the concerns described previously, this paper focusses on Indoor Positioning Systems (IPSs). These systems determine the location of the mobile node only when it is requested. For locating a mobile node, an IPS uses multiple anchor nodes. This is comparable with the principle of Global Positioning System (GPS) for outdoor, which uses satellites and a GPS-receiver.

In scientific literature, a large number of IPSs has been proposed. Unfortunately, most of these have been evaluated in non-healthcare related environments using only point accuracy. As already mentioned, for many healthcare use cases, in addition to point accuracy other relevant metric need to be taken into account. Each of which can influence the choice of the optimal technology. Moreover, these metrics may vary depending on a particular environment. In other words, an evaluation in an operational hospital environment is imperative to be able to asses real-life localization performances.

The main contributions of this paper are as follows. (i) A performance evaluation of multiple wireless IPSs is performed in an operational hospital environment that was actively in use and as such has a representative deployment of Wi-Fi Access Points (APs) and typical hospital interference. (ii) The impact of different design choices is quantified. The paper investigates the impact of the use of different localization algorithms, different wireless technologies and different anchor point locations. (iii) The performance of the different set-ups is evaluated using multiple evaluation criteria, including point accuracy, room accuracy and latency. (iv) The evaluation is focussed on stationary evaluation of localization solutions since the absence of history based location information is the most challenging. In this way, optimizations based on previous locations is excluded in the evaluation. (v) Finally, all data traces are made publicly available and can be used by third parties to evaluate additional IPSs.

The remainder of this paper is structured as follows. Section 4.2 gives an overview of IPSs with their classification and which ones are suitable for healthcare environments. This section also discusses other research papers that compare and evaluate multiple solutions and technologies. Section 4.3 describes the evaluation set-up, including the hospital environment, the used algorithms & hardware components and the evaluation methodology. Next, Section 4.4 discusses the performance evaluation for different set-ups and configurations. Finally, Section 4.5 concludes the paper.

4.2 Related Work

The introduction of this section is moved to Chapter 1, Section 1.1.2 on page 3 since this part gives the reader a general introduction about indoor localization solutions.

Typically, an IPS consists of an algorithm that processes wireless data from a specific technology. As such, an IPS can be seen as a combination of a localization algorithm running on top of a certain wireless hardware technology. Figure 4.1 presents the different layers. The main focus of this evaluation are the two lower layers: the technical performance. A localization algorithm can be classified in three categories as illustrated in Figure 4.2.

- The principle of **proximity algorithms** [6] is locating a mobile node using the highest received signal strength of an anchor node. The mobile node (which is accompanied with the object or person that needs to be located) is in the proximity of this anchor node whereof the highest signal strength signal came from. Typically, Near Field Communication (NFC) or Radio-Frequency Identification (RFID) is applied for this approach. Although Bluetooth Low Energy (BLE) is also capable to be used for proximity purposes. Proximity is easy to implement, does not require any complicated algorithms but the accuracy is low level, even room accuracy cannot be achieved. Since the accuracy is poor, this principle cannot meet the requirements from the hospital scenarios (described in the introduction).
- In contrast to proximity, **range based algorithms** use actual distances which are derived from the communication signals. A distinction between direction and distance based solutions can be made. Direction based means the direction of the propagation signal is the key element in determining the mobile node its position. Typically, an array of antennas or microphones is used to measure the angle between the signal and a reference. The spatial separation of antennas

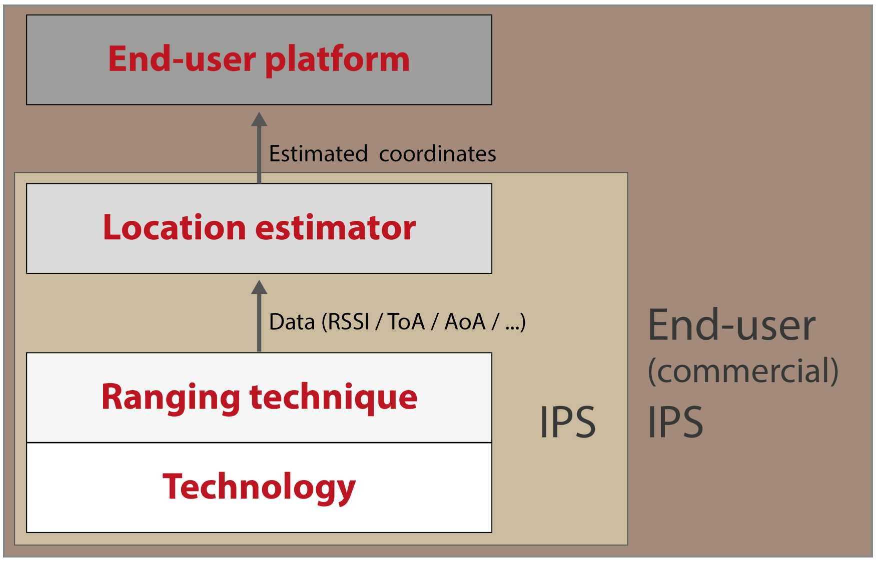


Figure 4.1: The four layers that define an end-user (commercial) IPS. The focus of this comparison are the three lower layers: the technology and ranging technique in combination with a certain location estimator. These three layers define an IPS whereby estimated coordinates of the mobile node are calculated.

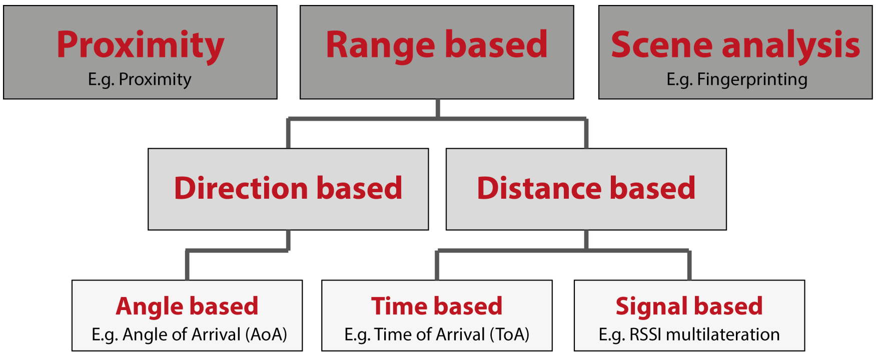


Figure 4.2: A classification of indoor localization solutions. Three categories are distinguished: proximity, range based and scene analysis. Range based can be split into direction or distance based. Direction based solutions uses the angle information of the antennas. Distance based solutions either use timing or signal based information.

or microphones leads to differences in arrival times, amplitudes and phases. The most typical example is the “Angle of Arrival (AoA)”-approach [7]. It can achieve a high accuracy, however it includes a significant hardware cost. Therefore, it is not implemented often in commercial applications.

Instead, algorithms based on the ranging distance are more popular. Two types can be differentiated: time or signal (property) based. Whilst time based algorithms (e.g. ToA [8] and Time-Difference of Arrival (TDoA)) determine distances based on the known signal propagation time, signal property based algorithms assume there is a proportional relationship between the Received Signal Strength Indicator (RSSI) and the distance. Generally, the main idea of range based algorithms remains the same: first, measured information (which may be derived from the angle, time or signal property information) is translated into a distance. Next, multiple distances are transformed into coordinates by applying Multilateration (MLAT).

- The final category in Figure 4.2 is **“Scene Analysis”**. The most typical example is Fingerprinting [9], which has a completely different approach than the ones described previously. This process is twofold. The first step (also training or offline step) includes an extensive survey of the environment whereby a set of training fingerprints (wireless characteristics, RSSI values of all available anchor nodes) is collected and stored into the training database. Second, the “online phase” consists of the location estimation. The currently measured wireless characteristics are compared with the fingerprinting database entries. The entry that matches the best will be used as the current location of the mobile node. Though this method of working is very accurate, it also has drawbacks. Completing this survey for an entire hospital is labor-intensive: every m^2 needs to be scanned and stored in a database. Even worse, environmental changes like moving a metal closet are impermissible and rescanning the environment is essential to keep the system accurate.

The goal of this paper is to identity which combinations of localization algorithms and wireless technologies are the most suited for hospital environments.

4.2.1 Comparison of Multiple Indoor Localization Solutions

To determine which solution is best suited, multiple relevant metrics have to be taken into account.

- **Room accuracy**: the possibility to locate a stationary mobile node at room level. E.g. locating an important but rarely used medical device can be equipped with a mobile node. If the position of this device is requested, room (and thus also floor) information can be sufficient.

- **Latency**: the time between sending a location request and receiving the location information. To continue with the previous example: the time it takes starting when a staff member sends a location request to locate a mobile node (which can be carried with a patient or attached to a medical device) until the staff member receives the location information that was requested. Another example is the "emergency call". When a patient pushed a mobile panic button, the latency of the localization solution can have an impact on the health status of the involved patient.

- **Installation time** / **cost**: hospital environments are (almost) continuously operational, meaning the installation time must be reduced to the minimum. Can the existing network be reused or is new wiring necessary? Does the solution requires recalibrating or not? Answers to those questions are reflected in the installation time and cost metric.

- **Energy consumption**: this metric is particularly important for the mobile node. This value is equivalent with the life-time of the device. A minimum duration of the mobile node can be required by the hospital.

The papers described below discuss multiple evaluation criteria, however these IPSs are not evaluated in an operational hospital environment.

In [10], a comparison of multiple Radio-Frequency (RF)-based indoor localization solutions in heterogeneous environments using multiple evaluation criteria is described. The authors conclude that the accuracy of the solutions depends strongly on the characteristics of the environment and that a fine grid of evaluation points is required for an objective comparison of solutions. Since the evaluation environments described in [10] consist of an office environment as well as an open industrial environment, this work motivates the need for extended evaluation testing in an operational healthcare environment.

Yanying Gu et al. compared indoor localization solutions with a special focus on the wireless personal networks [11]. In their comprehensive survey, they evaluate numerous solutions which include both commercial products and research-oriented solutions. Their evaluation criteria consists of security and privacy, cost, performance, robustness, complexity, user preferences, commercial availability and limitations. Their conclusions are in the same line as [12], each solution uses a certain type of technology, has its design and works well under certain situations.

[3] compares multiple IPSs to implement in a healthcare environment. However, they discuss the higher levels of integrating an IPS in a hospital

and their consequences: the impact and changes for the staff and patients, the Return On Investment (ROI) of an IPS, the possible risks when the system fails, etc. For those reasons, this paper is complementary to our work.

Finally, Sharif Vakili et al. compared a commercial and custom-made tracking solution [13]. Their comparison is comprehensive, using multiple evaluation criteria. Despite, these tracking solutions require manual actions from the users. Patients or nurses need to swipe a tag in front of a card-reader to indicate their entrance of the current room.

Taking into account the lessons learned from the related work, this paper will evaluate along multiple evaluation criteria for existing IPSs in an operational environment using a clearly defined methodology for objective evaluation.

4.2.2 Indoor Localization Solutions for Healthcare Environments

In scientific literature, several indoor localization solutions are proposed for the next generation of advanced healthcare applications [14–19]. However, they are not evaluated using the metrics above.

In [14], an indoor localization algorithm is described based on RSSI measurements that is optimized especially for the healthcare environments. Their solution guarantees room level accuracy while avoiding heavy investments by reusing the existing nurse call network. This approach achieves a high scalability since the mobile nodes locate themselves.

W. Chen describes in [15] a dynamic indoor localization solution based on active RFID. His algorithm is based on a cost function associated with a shape constraint factor. The cost function consists of the similarity and disparity of signal strength between the tracking and reference tags, as well as geometrical correlation properties. Results show that the proposed algorithm provides considerable improvement in average estimation error as compared with existing methods.

[16] presents an improved version of the low-frequency indoor localization system that is located under the floor. They achieved a larger detection range and a more durable antenna laminate. The measured tag detection was $2\ m$. The tag location reliability of $96.3\ \%$ was verified with a practical test.

In [17], a wireless localization network for patient tracking is presented. The network can track the locations of the patient and monitor their physical status i.e. walking, running, etc. by measuring their inertial movement using a three axis accelerometer. The Fleck-3 platform [18] is used for the static nodes. In this paper, a comparison is made between their own packet delivery ratio and the CC2431 Location Engine that used RSSI. This paper lacks any performance results like accuracy or latency and is only focussed on the network layer of the application.

A final example is LAURA [19], it stands for LocAlization and Ubiquitous monitoRing of pAtients for healthcare support. This solution is also using the signal strength of the ZigBee standard combined with a particle filter. LAURA achieves, both with static and moving patients, an average localization error lower than $2\ m$ in $80\ \%$ of the cases.

The mentioned papers above all describe a tracking solution designed and optimized for the healthcare sector. Some of them offer additional functionalities like patient monitoring. However, an objective evaluation approach is lacking. In many cases, no realistic hospital environment is used and multiple evaluation metrics like latency, installation cost, etc. are missing. This paper addresses these shortcomings.

4.3 Methods

The next section describes in details the hospital environment, the used hardware, the localization algorithms and finally the measurement execution.

4.3.1 Healthcare Environment

For the measurement campaign, an actively used hospital environment (the Sint-Jozefskliniek hospital in Izegem, Belgium) was selected. The measurements were performed in the "surgical day hospital" ward, located in a new building on the first floor. In this particular ward, patients arrive in the morning to undergo surgery and leave at night. The end section of the corridor was available to perform the experiments, while the rest of ward was in "normal operation", meaning patients and nurses were present and were walking around.

The floor plan of the ward is depicted in Figure 4.3. Rooms are located at both sides with "logistics" rooms in the middle. This means that there are two parallel corridors. Patient rooms 9, 10 and 11 were used for the evaluation. A dense evaluation grid of $1\ m$ by $1\ m$ was marked on the floor resulting in 73 evaluation locations where the position estimates were requested. Note that the grid was positioned in such a way that grid lines are $10\ cm$ away from the wall. During the experiment, all doors were open.

4.3.2 Installed Hardware

Anchor points from three different wireless technologies (Wi-Fi, ZigBee and BLE) were installed at the locations indicated on Figure 4.4. The locations are selected as realistic as possible. Wi-Fi APs are placed on the ceiling above each bed, whilst the ZigBee and BLE nodes are placed on the wall

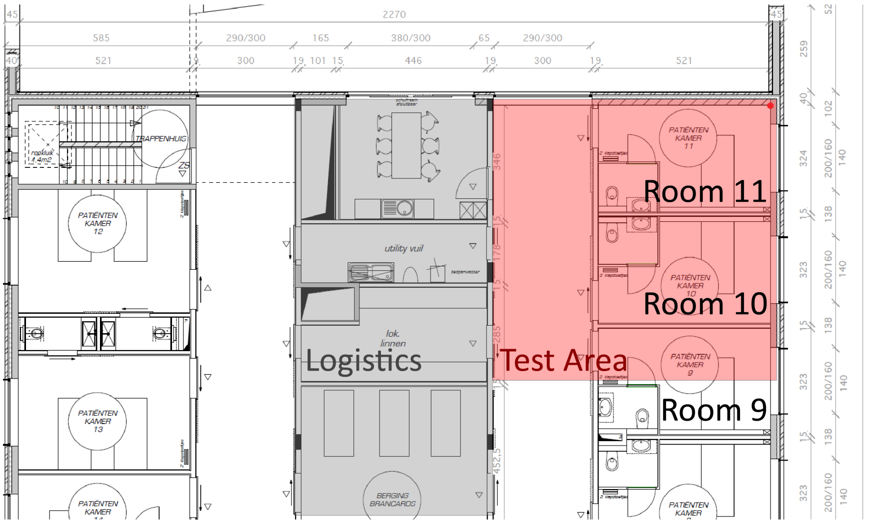


Figure 4.3: Floorplan of the hospital environment: rooms 9/10/11 and a hallway were used

nearby a light switch. Technical details of the devices can be found in Table 4.1.

Wi-Fi: A set of six Wi-Fi APs were deployed. One in each room in our test area. These APs are marked with a blue dot in Figure 4.4 (AP 30, 34, 28, 4, 90, 47).

ZigBee: A set of six ZigBee nodes (Zolertia Z1) were used during the measurement campaign. Their location is marked with a green dot in Figure 4.4 (1, 2, 3, 4, 5, 6).

BLE: Each sensor node was accompanied with a BLE beacon.

4.3.3 Localization Algorithms

During the evaluation, three different localization approaches were evaluated: a scene analysis algorithm (fingerprinting), a time based algorithm (ToA) and finally a signal based algorithm (MLAT). A detailed description of the algorithms can be found in [20]. Figure 4.2 shows how the algorithms can be classified according to the classification from Section 1.2 on page 5.

The fist solution is based on the fingerprinting principle. As mentioned in Section 1.2, this contains a twofold process whereby fingerprints are collected in a database during the learning phase. During the runtime phase, the current wireless statistics are compared and matched with the fingerprints in the database. In [21], the Wi-Fi network is used, but in theory

Figure 4.4: Floor plan of area in which the measurement campaign was performed. The evaluation points are located at the crossings of the orange grid lines. Deployed anchor points are indicated by blue dots (Wi-Fi access points) and green dots (ZigBee + BLE).

Table 4.1: Technical information of the setup in the hospital environment.

Anchor points	
Technology	**Technical details**
Wi-Fi	Netgear N750 Wireless Dual Band Gigabit Router
ZigBee	Zolertia Z1
BLE	BLE iBeacon (Estimote devices)
Mobile point	
Technology	**Technical details**
Wi-Fi	external 300Mbps Mini Wireless N USB adapter: TL-WN823N (TP-Link)
ZigBee	external STM32W-RCFKIT (using channel 25 and TX output power 31, 0dBm)
BLE	external Belkin mini Bluetooth v4.0 Adapter

any technology that contains RSSI values is possible. It is shown to be highly accurate, but it has drawbacks like installation and deployment time. This approach is also sensitive to changes in the environment. When this occurs, the training phase should be re-done.

A second approach is mainly based on the ToA-principle [22]. Time of Arrival localization solutions estimate distances between devices based on the propagation time of an RF wave between sender and receiver. Using the measured time and the speed of light, a corresponding distance can be determined. It is expected that the propagation time is linear correlated with the distance. The number of clock ticks is measured how long it takes to receive an acknowledgement when an unicast message was transmitted to a certain node. This approach is combined with a particle filter and is called “Spray”. Since it mainly uses ToA information, this approach can only be evaluated using ZigBee data.

Finally an RSSI MLAT based algorithm [12]. This approach is only based on the linear relationship between the RSSI value of the signal and distance between sender and receiver. Firstly, distance estimations of at least three different anchor nodes are retrieved during the ranging phase. In the second step, MLAT is applied in order to estimate the mobile node its position. Like the fingerprinting method, it only requires RSSI values and thus each technology is suitable.

4.3.4 Measurement execution

For the evaluation and comparison of different localization solutions the following approach is taken. (i) Packet transmitters (AP) of multiple technologies (Wi-Fi, ZigBee and BLE) are installed in an operational hospital environment. (ii) A fine evaluation grid consisting of evaluation points with known locations is established and drawn on the floor of the hospital. (iii) At each evaluation point packets from all APs are sampled. Since the measurement data is collected in an active hospital, with existing Wi-Fi access points as well as interference from other (medical) devices, realistic behaviour is obtained. (iv) During the data capturing phase, information traces from multiple technologies are annotated and stored separately. The data was captured during $30\ s$ with a laptop (moved around on a service cart) containing a dongle for each technology. An overview of the setup can be found in Figure 4.5. The technical specifications of the used dongles can be found in Table 4.1. (v) In order to capture the influence of the number and locations of access points, filters are applied on the datasets whereby one or more access points are removed so the robustness of an algorithm can be determined. (vi) Captured data is stored in the cloud and can be repeatedly used by a user to evaluate different algorithms. (vii) Once the System Under Test (SUT) produced a set of estimates, a set of metrics are calculated as follows. For all IPSs, the position error, room error and latency were calculated in the 73 evaluation locations and afterwards averaged. These metrics were calculated using the evaluation criteria from the EVARILOS benchmarking handbook [23] which is aligned with the upcoming ISO/IEC JTC 1/SC 31 standard for evaluating RF-based IPSs.

Both the raw datasets and the metric results are publicly available on the EVARILOS Benchmarking Platform (EBP). The EBP was already extensively used on multiple events (EOC [24], IPSN [25], etc.) and it was shown to be useful for objectively capturing the performance and comparing multiple solutions using multiple evaluation metrics.

4.4 Results and discussion

4.4.1 Impact of the Choice of the Algorithms

First, the performance results archived by different algorithms are compared to each other. For this evaluation, the data traces from all ZigBee node anchor points were given as input to all of the evaluated algorithms. ZigBee data is the only data source which may serve as input for all algorithms. The corresponding point accuracy is visualized in the form of heatmaps for each of the solutions in Figure 4.6. Blue areas refer to good accuracy results (point accuracy of $2\ m$ or less), whereas the accuracy worsens when the color changes to green, yellow and finally red. A red zone corresponds to a distance error around $10\ m$. A more detailed overview of the evaluation metrics using ZigBee can be found in Table 4.2.

Based on Figure 4.6, it is clear that fingerprinting approach achieves the most accurate results in general. The average error distance is $1.99\ m$. However, the latency is much higher than the one from other algorithms

Figure 4.5: The mobile node: a Dell laptop with 3 dongles (Wi-Fi, BLE, sensor node STM32) at a trolley at $100\ cm$ height. The ground truth of the evaluation points was indicated on the white sticker on the floor resulting in a gird of $1\ m$ by $1\ m$. Picture of a room and the corridor.

using the same data trace as input. In addition, fingerprinting solutions require a time-consuming calibration phase before they can be used, which might have to be repeated whenever the wireless environment changes significantly (for example due to the introduction of metal cupboards).

The spray solution is less accurate, achieving the average point accuracy around $3.89\ m$. In addition, in contrast to the previous solution, results show that the accuracy on one part of the environment is significantly higher than the accuracy in the other part. As a result, especially near the walls in the patient rooms, the corresponding room accuracy is significantly lower (Table 4.2).

Finally, the MLAT based approach is shown in Figure 4.6 (C). The average accuracy is around $4.06\ m$. It is clear that the additional deployment costs for calibrating fingerprinting based solutions results in significantly better accuracy results.

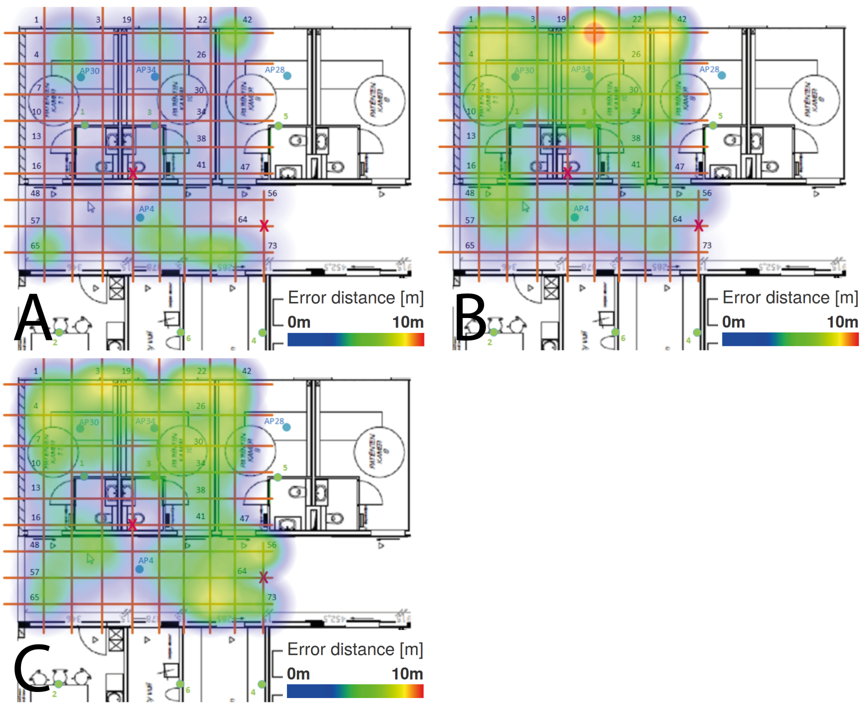


Figure 4.6: Heat maps representing the spatial distribution of localization errors of different localization algorithms using ZigBee data when all anchor nodes are used for location estimation. (A) Fingerprinting approach (B) Spray approach (C) MLAT approach

Table 4.2: Comparison of the evaluation metrics using the ZigBee dataset

Algorithm	Point Accuracy [m]	Room Accuracy [%]	Latency [s]
Fingerprinting	1.99	88	1.65
RSSI MLAT	4.06	49	0.50
Spray (RSSI + ToA)	3.89	47	0.50

4.4.2 Impact of the Choice of the Technology

Wireless technologies like Wi-Fi, BLE or ZigBee have the common possibility to retrieve a measure of signal strength during the wireless communication: RSSI. Since RSSI values are used as input for two out of the three evaluated algorithms, this fact allows us to investigate the influence of the wireless technology on the accuracy of a localization algorithm. The stability or the variance of RSSI values often depends on the technology, since different technologies have different methods for calculating RSSI and are impacted differently by interference. As described in [26], Wi-Fi suffers from the coexistence of BLE and ZigBee (and vice versa), since they all operate in the 2.4 GHz Industrial, Scientific and Medical (ISM) band.

The performance for each combination of technologies and algorithms is shown in Table 4.3. Since all algorithms use the same data trace as input, results can objectively be compared amongst different technologies and algorithms. Note that the spray algorithm requires time-of-arrival information, which is only available from the ZigBee nodes. As such, Spray is only evaluated using ZigBee data traces.

The results of fingerprinting are shown in Figure 4.7 (A) and (B). These are comparable with the results when ZigBee data was used (Figure 4.6 (A)). In general, this approach is very stable and achieve acceptable overall accuracy results. Further, no unexpected results are obtained for the MLAT approach (Figure 4.7 (C) and (D)). The error distances are comparable, except one outlier is detected when BLE data is used. These conclusions are reflected in the heat maps.

Table 4.3 shows that both solutions achieve the best accuracy results when Wi-Fi data is used. But at the same time Wi-Fi provides the worst latency results. One duty cycle for scanning for available networks takes $3\ s$. This cannot be interrupted.

As a conclusion, the differences between the technologies are minimal. Wi-Fi is slightly better and similar performance was achieved by ZigBee and BLE. A possible explanation is the difference in output power of these technologies. Wi-Fi's output power is higher whilst the output power of ZigBee and BLE is quite similar.

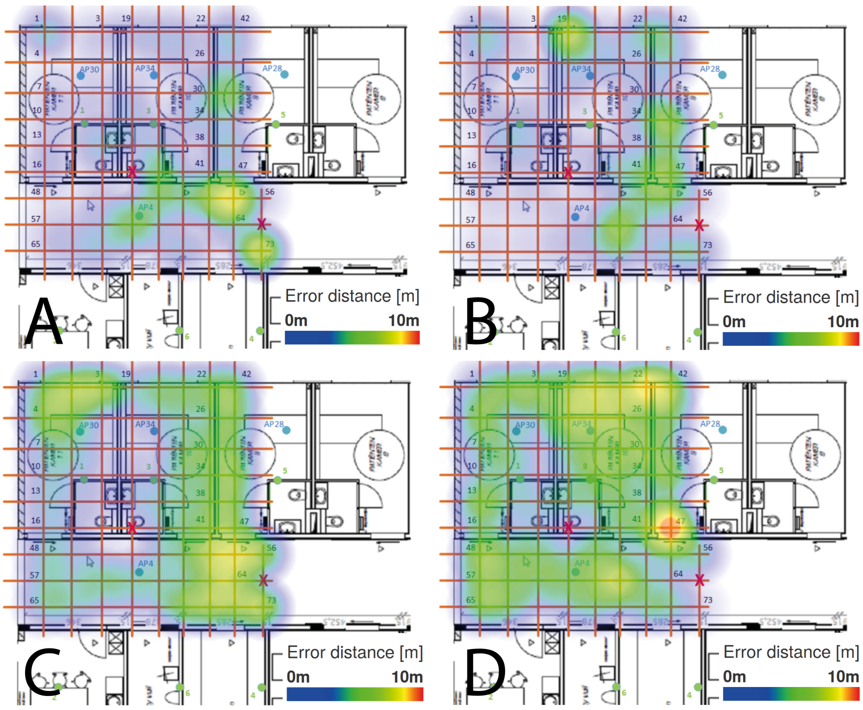


Figure 4.7: Comparison of point accuracy of different technologies and algorithms. (A) Fingerprinting, Wi-Fi (B) Fingerprinting, BLE (C) MLAT, Wi-Fi (D) MLAT, BLE

Table 4.3: Comparison of the evaluation metrics using the full dataset

Technology	Point Accuracy [m]	Room Accuracy [%]	Latency [s]
	Fingerprinting		
ZigBee	1.99	88	1.65
Wi-Fi	1.21	96	5.43
BLE	2.13	79	3.06
	RSSI MLAT		
ZigBee	4.06	49	0.50
Wi-Fi	3.65	47	3.00
BLE	3.85	61	2.50
	Spray RSSI + ToA		
ZigBee	3.89	47	0.50

4.4.3 Impact of the Choice of the Anchor Point Selection

A final analysis will be discussed in this subsection: the influence of the available anchor points. Anchor points have a huge impact on installation time and robustness of the solution (in case an anchor point fails). To limit the number of heat maps, only the influence of access points with the MLAT algorithm is discussed.

In the previous sections, all available access points were used. The same datasets of the previous sections could be reused since additional filter techniques are applied. In this way, a perfect comparison is possible, since datasets contain the same interference and pedestrian pattern.

Figure 4.8 shows all the heat maps (of each technology) whereby a different set of anchor points is used. Initially all six anchor nodes were used, the location of these anchors can be found on the map in Figure 4.4. For each technology, two different subsets are created. For Wi-Fi, the first one is without the centre AP located in the corridor and AP 4 (Figure 4.8 (A)). In this situation, the algorithm still achieves stable results. The changes are minimal compared with Figure 4.7 (C)). The error distances increase when only one side of the corridor is equipped with APs. In this case, AP 28, 30 & 34 are used. The solution performs weak mainly in the corridor and the third patient room. The point accuracy results remain more or less stable. On the other hand, the room accuracy drops drastically, from 56% to 31% (Table 4.4).

The subset “Only edges” contains the anchors located in the corners of the evaluation environment (1, 2, 4 & 5). If BLE is the used technology, the algorithm has high error distances at the east side of the corridor (Figure 4.8 (C)). A big contrast when all anchor points are available (Figure 4.7 (D)). It is even worse when ZigBee data is used (Figure 4.8 (E)). The solution obtains a few error distances around $10\ m$ at the east side of the corridor, but the worst results are retrieved at the top of the test area, in a patient room.

Figure 4.8 (D) shows the spatial distribution of the localization error when using only the anchor points at the same side of the corridor (Anchor 1, 3 & 5). Higher errors can be observed in the corridor. In the patient rooms the algorithm works well as the evaluation points are in close proximity while in the corridor this is not the case. Having anchors on one line is clearly not a good option. For ZigBee, the same conclusion holds (Figure 4.8 (F)).

Anchor nodes have a strong influence on the MLAT algorithm. Explicable, since the approach is based on MLAT, results are worse if the anchors are positioned in one line instead of a triangle [27]. In comparison, fingerprinting remains more stable when certain anchor nodes are unavailable.

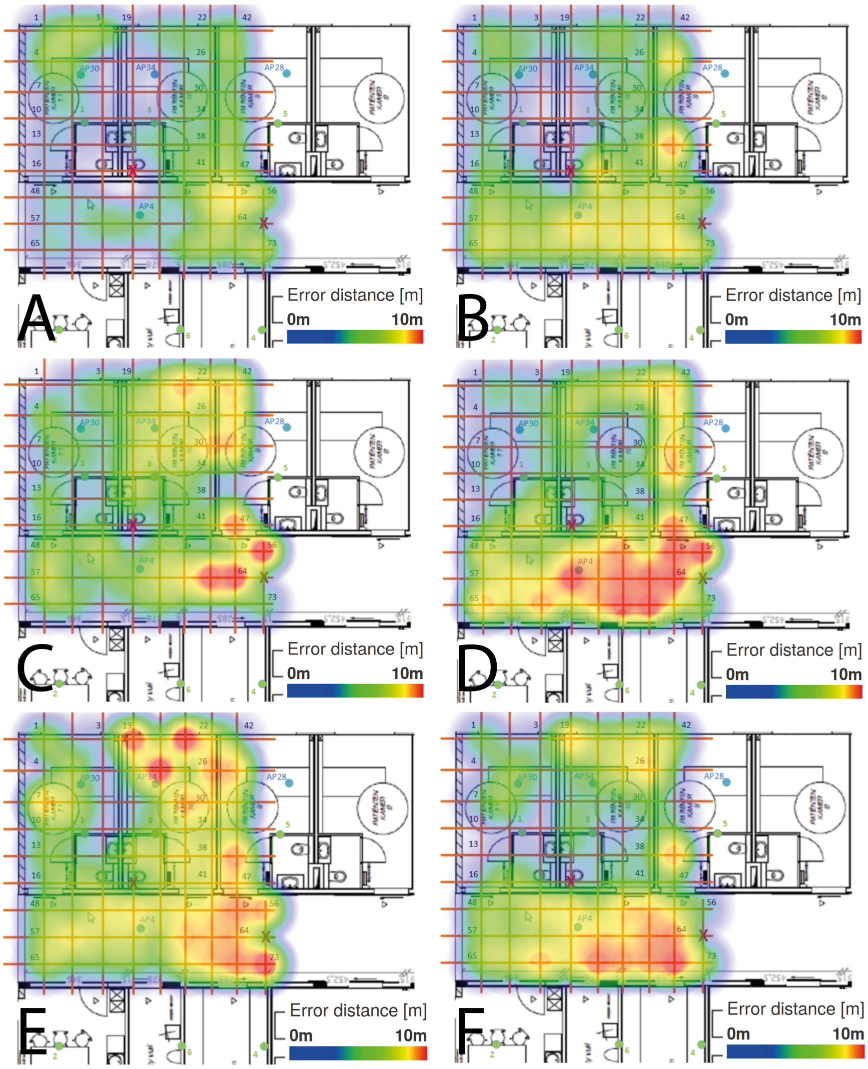


Figure 4.8: Heat maps representing the point accuracy of MLAT using all kind of data with different amount of anchor nodes. (A) Wi-Fi, without AP 4 (B) Wi-Fi, one side corridor (C) BLE, only edges (D) BLE, one side corridor (E) ZigBee, only edges (F) ZigBee, one side corridor

Table 4.4: Accuracy results with different subsets of anchor points

Filter	Point Accuracy [m]			Room Accuracy [%]
	Minimum	Average	Maximum	
Wi-Fi				
All anchors	0.52	2.68	5.95	46.58
Without AP4	0.07	2.69	6.77	56.16
One side corridor	0.48	3.47	8.57	31.51
BLE				
All anchors	0.94	3.12	9.09	61.64
Only edges	0.54	4.04	10.49	41.10
One side corridor	0.48	4.78	10.49	36.99
ZigBee				
All anchors	0.10	3.08	7.35	49.32
Only edges	0.48	5.09	10.89	21.92
One side corridor	0.48	4.24	9.94	38.36

4.4.4 Discussion

When accuracy is the most deciding factor, fingerprinting lends itself to be the solution that should be implemented in combination with Wi-Fi access points. Moreover, fingerprinting proofs to be the most robust as well. When one of the APs fails, it will preserve its accuracy results. However when other parameters like installation time, “environment robustness” and latency have an influence on the decision, other approaches like Spray might become interesting.

4.5 Conclusion

The need for more advanced systems is rising in the healthcare industry. Nurse calling or patient tracking systems need accurate and always up to date location information. Once this information is adopted in the previous mentioned systems, a whole of new services will become available. Therefore, in this paper, a thorough analysis of multiple facets of indoor localization approaches in an healthcare environment is executed.

Firstly, multiple algorithms are evaluated using the same amount and type of data. A fingerprinting, ToA and MLAT approach are compared using one single dataset, recorded in an operational hospital environment. Based on the accuracy results, fingerprinting achieves the best score ($1.21\ m$).

But on the other hand, fingerprinting has the highest latency and the worst installation and configuration time. A trade-off must be made depending on the primary requirements.

A second validation was the type of technology. Wi-Fi, BLE and ZigBee data was recorded during the measurement campaign in the hospital. The influence of the technology seems to be minimal on the accuracy metrics. However, latency impact is different: Wi-Fi has a duty cycle of three seconds. It takes at least three seconds before any RSSI data is available. BLE and ZigBee their update cycles are much shorter. In conclusion, when accuracy matters the most, Wi-Fi technology along with a fingerprinting algorithm yields the best result. The reason for that can be found in a higher bandwidth and transmission power of Wi-Fi in comparison to other technologies, which results in more stable RSSI measurements and higher coverage of Wi-Fi signals in the environment. When latency, power consumption or deployment matter the most, a "cheap" technology, such as BLE or ZigBee, are a decent alternative. When you can realize a dense deployment easily (at least one per room), the accuracy can be very good as well but more nodes are needed than in the Wi-Fi case.

Finally, an influence of the anchor nodes was evaluated. It is crucial to know how many anchor nodes are required for achieving accurate results and how a failure of one single anchor node can influence the stability of the entire system. Two different subsets are compared with the original situation. For the MLAT approach, the impact is of missing anchors is clearly visible at the heat maps. In case Wi-Fi is used, the area where the mobile node is localized, should be as much as possible within the anchor points. Optimally the resolution in X and Y directions is similar. This means that access points in every direction surround you. Additional anchors in the corridor are not required as they do not significantly improve the accuracy. Typically, one access point per every two rooms is a good compromise between accuracy and deployments. If ZigBee or BLE is used, a denser deployment is required than in the case of Wi-Fi. Of course these nodes are cheaper and consume less energy. A node per room is required. Nodes should be present in the rooms at both sides of the corridor. In that case, no additional nodes in the corridor are needed.

In general, Wi-Fi technology has most potential for cases where accuracy matters the most. The complexity of the algorithm is more important than the raw technology choice. ZigBee and BLE technologies show very similar results. A Wi-Fi fingerprinting solution with an anchor installed in every two rooms would be the preferred solution for a hospital environment.

References

[1] J. Buchan and L. Aiken. *Solving nursing shortages: a common priority*. Journal of clinical nursing, 17(24):3262–3268, 2008.

[2] W. C. Mann. *The aging population and its needs*. Pervasive Computing, IEEE, 3(2):12–14, 2004.

[3] M. N. K. Boulos and G. Berry. *Real-time locating systems (RTLS) in healthcare: a condensed primer*. International journal of health geographics, 11(1):25, 2012.

[4] A. L. Tucker and S. J. Spear. *Operational failures and interruptions in hospital nursing*. Health services research, 41(3p1):643–662, 2006.

[5] N. M. Drawil, H. M. Amar, O. Basir, et al. *GPS localization accuracy classification: A context-based approach*. Intelligent Transportation Systems, IEEE Transactions on, 14(1):262–273, 2013.

[6] X. Jiang, C.-J. M. Liang, K. Chen, B. Zhang, J. Hsu, J. Liu, B. Cao, and F. Zhao. *Design and evaluation of a wireless magnetic-based proximity detection platform for indoor applications*. In Proceedings of the 11th international conference on Information Processing in Sensor Networks, pages 221–232. ACM, 2012.

[7] P. Kułakowski, J. Vales-Alonso, E. Egea-López, W. Ludwin, and J. García-Haro. *Angle-of-arrival localization based on antenna arrays for wireless sensor networks*. Computers & Electrical Engineering, 36(6):1181–1186, 2010.

[8] Y.-T. Chan, W.-Y. Tsui, H.-C. So, and P.-c. Ching. *Time-of-arrival based localization under NLOS conditions*. Vehicular Technology, IEEE Transactions on, 55(1):17–24, 2006.

[9] W. Meng, W. Xiao, W. Ni, and L. Xie. *Secure and robust Wi-Fi fingerprinting indoor localization*. In Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on, pages 1–7. IEEE, 2011.

[10] T. Van Haute, E. De Poorter, I. Moerman, F. Lemic, V. Handziski, W. Adam, N. Wirström, and T. Voigt. *Comparability of RF-based Indoor Localization Solutions in Heterogeneous Environments: An Experimental Study*. International Journal of Ad Hoc and Ubiquitous Computing, 1(1):1–23, 2015.

[11] Y. Gu, A. Lo, and I. Niemegeers. *A survey of indoor positioning systems for wireless personal networks*. Communications Surveys & Tutorials, IEEE, 11(1):13–32, 2009.

[12] T. Van Haute, J. Rossey, P. Becue, E. De Poorter, I. Moerman, and P. Demeester. *A hybrid indoor localization solution using a generic architectural framework for sparse distributed wireless sensor networks*. In Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on, pages 1009–1015. IEEE, 2014.

[13] S. Vakili, R. Pandit, E. L. Singman, J. Appelbaum, and M. V. Boland. *A comparison of commercial and custom-made electronic tracking systems to measure patient flow through an ambulatory clinic*. International journal of health geographics, 14(1):1, 2015.

[14] J. Wyffels, J. De Brabanter, P. Crombez, P. Verhoeve, B. Nauwelaers, and L. De Strycker. *Distributed, Signal Strength-Based Indoor Localization Algorithm for Use in Healthcare Environments*. Biomedical and Health Informatics, IEEE Journal of, 18(6):1887–1893, 2014.

[15] W.-H. Chen, H. H. Chang, T.-H. Lin, P.-C. Chen, L. Chen, S. Hwan, D. Yen, H. S. Yuan, and W. C. Chu. *Dynamic indoor localization based on active RFID for healthcare applications: a shape constraint approach*. In Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on, pages 1–5. IEEE, 2009.

[16] A. Ropponen, H. Rimminen, and R. Sepponen. *Robust system for indoor localisation and identification for the health care environment*. Wireless Personal Communications, 59(1):57–71, 2011.

[17] M. D'Souza, T. Wark, and M. Ros. *Wireless localisation network for patient tracking*. In Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference on, pages 79–84. IEEE, 2008.

[18] P. Sitka, P. Corke, L. Overs, P. Valencia, and T. Wark. *Fleck-a platform for real-world outdoor sensor networks*. In Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on, pages 709–714. IEEE, 2007.

[19] A. Redondi, M. Tagliasacchi, M. Cesana, L. Borsani, P. Tarrío, and F. Salice. *LAURALocAlization and Ubiquitous monitoRing of pAtients for health care support*. In Personal, indoor and mobile radio communications workshops (PIMRC Workshops), 2010 IEEE 21st International Symposium on, pages 218–222. IEEE, 2010.

[20] P. Crombez, J. Gesquiere, G. Glorioso, C. Villoute, W. Yves, N. Wirström, T. Van Haute, E. De Poorter, V. Handziski, F. Lemic, and A. Wolisz. *D4.2 Report on the Results of the Real-life Experiments in the Validation Scenarios*. EVARILOS project.

[21] F. Lemic. *Benchmarking of Quantile-based Indoor Fingerprinting Algorithm*. Telecommunication Networks Group, Technische Universität Berlin, Tech. Rep. TKN-14-001, 2014.

[22] N. Wirstrom, P. Misra, and T. Voigt. *Spray: A multi-modal localization system for stationary sensor network deployment*. In Wireless On-demand Network Systems and Services (WONS), 2014 11th Annual Conference on, pages 25–32. IEEE, 2014.

[23] T. Van Haute, E. De Poorter, F. Lemic, V. Handziski, T. Voigt, A. Wolisz, I. Moerman, et al. *Platform for benchmarking of RF-based indoor localization solutions*. Communications Magazine, IEEE, 53(9):126–133, 2015.

[24] F. Lemic, V. Handziski, A. Wolisz, T. Constambeys, C. Laoudias, S. Adler, S. Schmitt, and Y. Yang. *Experimental Evaluation of RF-based Indoor Localization Algorithms Under RF Interference*. Proc. ICL-GNSS'15.

[25] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen. *A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned*. In Proceedings of the 14th International Conference on Information Processing in Sensor Networks, pages 178–189. IPSN '15, 2015.

[26] L. Pei, J. Liu, R. Guinness, Y. Chen, T. Kroger, R. Chen, and L. Chen. *The evaluation of WiFi positioning in a Bluetooth and WiFi coexistence environment*. In Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS), 2012, pages 1–6. IEEE, 2012.

[27] H. S. AbdelSalam and S. Olariu. *Towards enhanced RSSI-based distance measurements and localization in WSNs*. In INFOCOM Workshops 2009, IEEE, pages 1–2. IEEE, 2009.

Part III

Guidelines for Benchmarking Indoor Localization Solutions

5

The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions

Part III bundles guidelines and tools for benchmarking indoor localization solutions, it addresses the shortcomings which are identified in Chapter 3. This methodology is verified and optimized in Chapter 4. The methodology includes definitions of metrics and benchmarking scenarios. These scenarios define the necessary steps to evaluate an Indoor Positioning System (IPS). The most recent version of this handbook and its methodology can be found at the EVARILOS website[1].

⋆ ⋆ ⋆

T. Van Haute, E. De Poorter, J. Rossey, I. Moerman, V. Handziski, A. Behboodi, F. Lemic, A. Wolisz, N. Wirström, T. Voigt, P. Crombez, P. Verhoeve, J. de las Heras.

Published in the proceedings of MERMAT, April 2013.

[1] http://www.evarilos.eu/

Abstract RF-based indoor localization solutions enjoy consistent efforts of researchers to provide more accurate and sustainable solutions. The multiplicity of RF-based indoor localization solutions makes their evaluation an indispensable part of future Internet. However no unified scheme has been devised for evaluation of these solutions and their robustness against various parameters. To remedy this, the EVARILOS handbook is created in order to objectively evaluate and compare different indoor localization solutions. In this work, we present an overview of the EVARILOS project whose objectives are the development and validation of standardized experiment-based benchmarks for localization solutions.

5.1 Introduction

Accurate and robust indoor localization is a key enabler for context-aware Future Internet applications, whereby robust means that the localization solution should perform well in diverse physical indoor environments under realistic RF interference conditions. However, despite the abundance of works on RF-based indoor localization solutions, the numerous published solutions are evaluated under individual, not comparable, not repeatable and often not realistic conditions. No unified scheme is provided for the fair comparison and evaluation of various solutions. Therefore it is necessary to develop and establish a comprehensive benchmarking methodology which is able to consider variety of existing solutions and their significant features. The EVARILOS project (**Eva**luation of **R**F-based **I**ndoor **Lo**calization **S**olutions for the Future Internet) [2] focuses on the development of the benchmarking methodology which consists of providing (i) metrics for evaluation of RF-based indoor localization solutions and (ii) a set of benchmarks and scenarios which are recommended to use for experimental performance evaluation according to the previous metrics for a given solution.

The main outcomes of the project are a public handbook on the use of the EVARILOS benchmarking methodology and the EVARILOS benchmarking suite. The benchmarking suite will be publicly available under open source licenses and implemented in two different testbeds belonging to the FIRE facilities (FP7 CREW [3] and FP7 OpenLab [4]), more specifically on the testbeds in Berlin and Ghent. The EVARILOS project uses the OMF [5] control and management framework and mobility support features developed in OpenLab and will further use and extend the benchmarking features from CREW. An open challenge is also envisaged using the above mentioned testbeds to invite external experimenters for evaluation of their

[2] http://www.evarilos.eu
[3] http://www.crew-project.eu
[4] http://www.ict-openlab.eu
[5] http://mytestbed.net

localization solutions and use their feedback and results together with the results of our own experiments to create the first repository of localization solutions evaluated using a unified methodology.

The rest of this paper is organized as follows. The next section handles about the state-of-the-art of indoor localization. The benchmarking methodology in Section 5.3 describes the general structure of the benchmarking handbook. The scenarios, environment and metrics are further explained in respectively Sections 5.4, 5.5 and 5.6. In Section 5.7, the future work of the EVARILOS project is described. Finally, conclusions are made in Section 5.8.

5.2 Related work

Generally, there are two phases towards realization of accurate location-based applications: **ranging** and **location estimation**. A state-of-the-art overview is given of existing ranging techniques and location estimation methods.

5.2.1 Ranging

Localization methods can be divided into two categories [1] *range-based* and *range-free*. The former is defined by protocols that use absolute point-to-point distance estimates (range) or angle estimates for calculating location. The latter makes no assumption about the availability or validity of such distance or angle information. Range-free methods are found in rather theoretical, not empirical work. [2] compares two range-free localization algorithms. In environments with obstacles, many range-free techniques that have been proposed to improve the localization accuracy are useless and inversely decrease the algorithm's accuracy [3]. The most commonly used techniques to perform ranging are:

RSSI (Received Signal Strength Indication) is an indication of the power level received by a receiver expressed in dBm. This value is then used to estimate the distance between transmitter and receiver. The physics behind this technology is the power level decay with distance. RSSI is available in most RF receivers.

ToA (Time of Arrival), also called ToF (Time of Flight), uses the travel time of a radio frequency wave from one transmitter to one receiver. With the speed of light the distance is calculated. ToA requires precise synchronization of timers at both transmitter and receiver.

TDoA (Time Difference of Arrival) is also based on the speed of light. Here the position is calculated with at least three spatially separated receiver sites (and one transmitter, being the object to be localized).

The difference of the time of arrival at two receivers will narrow the possible position to one half of a twosheeted hyperboloid. The knowledge of the time of arrival at the third receiver is needed to calculate the unknown position. TDoA only requires precise synchronization at the receivers. In many wireless sensor networks, TDoA is based on the time difference between simultaneously transmitted radio and ultrasound pulses as in the Cricket system, as typical WSN clocks are too slow for the first approach only [4].

AoA (Angle of Arrival) determines the angle of an incident RF wave, which requires special antennas such as antenna arrays. AoA methods based on antenna arrays determine the direction by measuring the time difference of arrival (delays) at individual antenna elements of the array. From the delay measurements at the individual antenna elements, the angle of arrival can be calculated. Because most antennas are reciprocal, this can be considered as reverse beamforming.

DTDoA (Differential Time Differences of Arrival) uses the difference of TDoA measurements. This is done to overcome the time synchronization of both transmitter and receivers. This is accomplished by introducing a fourth anchor that is responsible for initiating the TDoA measurement by transmitting a special message. In this way the anchors' time offsets can be computed [5].

Proximity uses a very weak sending power, if a message is received, the receiver knows it is the vicinity of the sender. We cannot infer anything, if the message is not received.

Hybrid techniques are the combination of any of the previous techniques.

An overview of the different ranging techniques with the different wireless technologies can be found in Table 5.1.

5.2.2 Location estimation

Once the ranging measurements are available between the fixed anchor points (whose position is already known) and the (mobile) object to be located (whose position is unknown), it is possible to utilize several methods for the estimation of the location of the object.

A first distinction can be made between fingerprinting or not, typical for a fingerprinting is the use of a large database and training phase. This database is filled with measurements (e.g. the RSSI-values recorded by nodes knowing their own position) during the time consuming (in the order of several days) off-line phase (also called training phase). The online phase is the positioning of a target: here a new measurement is compared to the values in the database. The stored measurement that is closest to

Table 5.1: Wireless technologies for indoor localization versus ranging techniques

	Ultrasonic	Infrared	IEEE 802.15.1	RFID	IEEE 802.11	IEEE 802.15.4a DSSS	IEEE 802.15.4a UWB	IEEE 802.15.4a CSS	60 GHz
RSSI	✗	✗	✓	✓	✓	✓	✓	✗	✗
ToA (ToF)	✓	✗	✗	✓	✓	✓	✓	✓	✗
TDoA	✓	✗	✗	✓	✓	✗	✓	✗	✗
DTDoA	✗	✗	✓	✗	✓	✗	✗	✗	✓
AoA	✓	✗	✗	✓	✓	✓	✓	✗	✗
Proximity	✓	✓	✓	✓	✓	✓	✗	✗	✗

the measurement of the target gives the estimated position. A drawback of this method is that the database needs to be filled with new measurements if the environment changes (e.g. adding a new bed in a hospital localization system [6, 7]). Non-fingerprinting methods do not require an off-line phase and these methods are faster.

A distinction is based on the usage of geometric techniques or statistical methods. Geometric techniques use geometry to calculate the position from at least three ranging measurements. An example is geometric multilateration [8]. The use of statistics is widely accepted in location estimation. Here, the frequency distribution of the distances is considered for making an estimation of the position. Mainly, there are three different methods: statistical multilateration, maximum likelihood estimators and Bayesian estimators.

In its simplest form statistical multilateration [8] minimizes the sum of the squares of the ranging errors (e.g. distance errors). There also exists a weighted least square approach, like in [9]. Here the measured values are first weighted, before the minimization: e.g. high RSSI-values are given a higher weight. In indoor environments this leads to the unjust preference of the paths with the most constructive multipath fading according to a study recently performed by iMinds [10].

Maximum likelihood methods make use of a cost function. Dependent on the kind of cost function, it needs to be minimized or maximized to find the most likely position. Several cost functions exist. In [11] the simplest and most widely accepted method (minimum mean square error) is presented. Some more cost functions, not only for RSSI but also for ToA measurements, are presented in [12]. A linear regression based cost function

Figure 5.1: The schematic representation whereby measurements are transformed into scores using metrics. In this example, a distance measurement is translated into a score (a distance) using a meter.

is introduced by iMinds in [13].

Bayesian localization methods are based on Bayes' theorem [14] and therefore incorporate some prior knowledge in the estimator. Two extensively used methods are Kalman filters, Particle filters [15, 16] and hidden Markov models (HMM) [17]. Another example of a Bayesian method can be found in [18].

5.3 Benchmarking methodology

As stated earlier, the EVARILOS project addresses one of the major problems of indoor localization: the pitfall to reproduce research results in real life scenarios and the inability to compare their performance due to evaluation under individual, not comparable and not repeatable conditions. The EVARILOS handbook presents a benchmarking methodology that remedies these shortcomings, by defining objective experimental validation of and fair comparison between state-of-the-art indoor localization solutions under different use-case scenarios and configuration setups.

Contrary to previous approaches, the EVARILOS benchmarking methodology does not focus exclusively on the accuracy of the evaluated localization approach, but also considers other important criteria that are relevant in view of the commercial deployment of localization solutions such as complexity, interference robustness, cost, energy efficiency, etc. Since different use cases have different sensitivities for individual metrics, the EVARILOS benchmarking methodology cleanly decouples between the metrics and the calculation of the final score used for ranking. As illustrated on Figure 5.1, after collecting a set of measurements necessary for the calculation of the individual metrics, the EVARILOS methodology allows application of specific weighting factors for the calculation of the final ranking score that reflect the different impact of the metrics for the different application scenarios of interest.

In the EVARILOS point of view, a benchmark is an evaluation method that is used to evaluate and compare the performance of one or more localization solutions. A benchmark is a combination of environment specifications, the setup and unambiguously defined performance metrics. An EVARILOS benchmark allows a fair and objective comparison of different localization solutions such that they can be ordered by a binary relation $\leq$. The object of comparison is the benchmark score(s).

5.4 Scenarios

A scenario fully describes a benchmark, and consists of a definition of the used metrics, the criteria of the evaluation and all the necessary parameters and traces to perform the experiment. A scenario description is a combination of an *environment description*, a *setup description* and specification of the *metrics*. This is illustrated in Figure 5.2. All evaluations are considered black box benchmarks: the scenario description can be seen as a black box, which takes as input a localization approach, and outputs one or more numerical benchmarking values. As such, the internal properties of the localization scenario are not evaluated, only its relevance for different application domains.

The next two sections elaborate on the environment specifications and the evaluation metrics of a scenario.

5.5 Environment description

An environment specification is the description of the physical environment and the infrastructure that is used to perform an experiment. An environment is typically represents a real-life situation, for example an office environment. As such, the environment description defines both structural properties of the environment (e.g. room layout, room sizes, types of walls, etc.) and RF interference properties (e.g. what types of external RF technologies are present and to what extent). The performance of a localization solution is always related to a specific environment. The environment consists of two parts: the building specifications and the interference specifications of the environment.

5.5.1 Building specifications

Building specifications represent the infrastructure of a specific environment. In the benchmarking handbook, three types of walls are determined: open space (no walls), (ply)wooden walls and brick walls. For each type of wall, a corresponding room size must be selected (Small, medium or

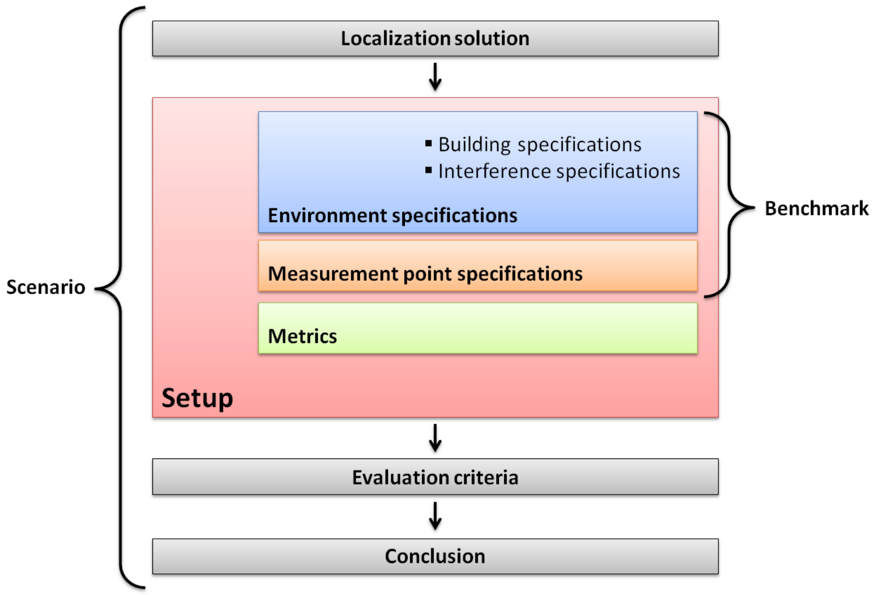


Figure 5.2: Visual representation of a scenario and its components. The input is a localization solution. After specifying the setup and the evaluation criteria, some conclusion can be made.

big). Since the performance of a localization solution is often strongly related to the type of environment, all benchmarking outputs must always be given together with a description of the building specifications. For fair comparison, the handbook describes in detail a number of predetermined reference building types.

5.5.2 Interference specifications

The list of interference specifications is more complex than the building specifications. Four different types can be distinguished: no, low, moderate and high interference. There are many parameters that define a certain interference profile. (i) Network parameters, e.g. network size, node density, mobility or failures, etc. (ii) Traffic parameters, e.g. packet size, inter packet gap, bitrate, filesize, start & stop time, traffic model, etc. (iii) Parameters of the interference source, e.g. number of sources, power, waveform, pattern, etc. and finally (iv) different types of interference, e.g. microwave, Wi-Fi, Bluetooth, 3G, Zigbee, etc. Interference can be created artificially, or by replaying previously captured interference traces. Again, for fair comparison, the handbook describes a number of predetermined reference interferences types.

5.6 Metrics

A metric is a measure of a specific performance indicator of the system under test. Depending on the type of metrics, for example accuracy, installation costs, etc., metrics are classified as deployment, functional or performance metrics. For comparing the suitability of a solution for a specific application domain, weight factors are assigned to the different metrics.

The EVARILOS benchmarking methodology takes into account the multifaceted nature of localization schemes and strives to define an adequate ensemble of metrics for evaluation process. For each individual metric, a definition is given, together with instructions for collecting the necessary underlying measurements and a mathematical formula that should be used for processing those measurements in order to calculate the metric value. The metrics that should be calculated depend on the application scenario, that describes which metrics are required, and which weighting factors are used for the calculation of the final ranking score. For each metric of interest, the handbook then recommends a set of benchmarks for the experimental assessment of the performance. The metrics are organized in three generic categories: performance metrics, deployment metrics and functional metrics. A structural overview is given in Figure 5.3.

The first and largest category is comprised by several metrics that try to capture different performance aspects of the system under test, such as its accuracy, robustness, scalability, etc. In this category, a distinction between the primary performance (Subsection 5.6.1 and 5.6.2) and derived performance metrics (Subsection 5.6.3 and 5.6.4) is made. The latter can be measured using the primary performance metrics. These, so called derived performance metrics, represent the sensitivity of the solution to different (external) factors, such as interference or mobility speed.

To calculate them, the accuracy is first measured in simple controlled environments before determining the sensitivity to external changing conditions. The functional metrics focus on non-performance related attributes like the underlying technology, licensing modalities, open-source availability, etc. Finally, the deployment metrics capture important properties related to the efforts and costs needed for physical installation, configuration, and replacement time.

5.6.1 Accuracy

In the EVARILOS benchmarking, two different accuracy metrics are usually used: point and room accuracy. With point accuracy, the actual Euclidean error distance between a reference point and a measured point is calculated. Suppose the reference point has coordinates (x_1, y_1, z_1) and the measured point (x_2, y_2, z_2), than the error distance d can be found by using Equation 5.1 for a 2D and Equation 5.2 for a 3D coordinate system.

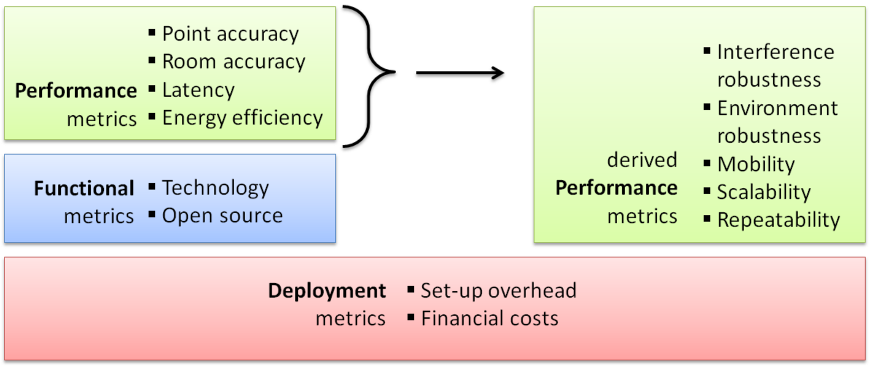


Figure 5.3: The tree structure of the metrics. Four different groups can be differentiated: performance, functional, deployment and derived performance metrics.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{5.1}$$

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{5.2}$$

Once the distances of the multiple tests are calculated the mean, standard deviation, minimum and maximum values can be calculated using the following equations:

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i \tag{5.3a}$$

$$\sigma_d = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i - \overline{d})^2} \tag{5.3b}$$

$$d_{min} = \min(d_1, d_2, ..., d_n) \tag{5.3c}$$

$$d_{max} = \max(d_1, d_2, ..., d_n) \tag{5.3d}$$

For some applications, room accuracy is required instead of point accuracy. When room accuracy is required, the localization solution only needs to identify the room in which the node is located. A distinction is also made between different floor levels.

To visualize the results of the room accuracy, a room confusion matrix is used. Each column of the matrix represents the instances in a predicted room, while each row represents the instances in an actual room. An example of a confusion matrix is given in Table 5.2 with the assumption that each room is located next to each other and each room is tested 10 times.

Table 5.2: A confusion matrix: example

		Predicted room				
		Room 1	Room 2	Room 3	Room 4	Room 5
Actual room	Room 1	**7**	2	1	0	0
	Room 2	1	**8**	1	0	0
	Room 3	1	2	**6**	0	1
	Room 4	0	1	0	**9**	0
	Room 5	0	0	2	1	**7**

In Table 5.2 the number of correct rooms is in bold (the predicted room corresponds to the actual room). The other numbers are the amount of incorrectly predicted rooms. With these numbers, a simple success rate can be calculated by dividing the number of correct rooms by the total number of rooms available. This becomes clear in Equation 5.4. Even more sophisticated success rate equations can be used where the geographical position of the rooms can be taken into account.

$$sr = \frac{\textit{number of correct rooms}}{\textit{total number of rooms}} \tag{5.4}$$

Both for point and room accuracy, a localization path is constructed that is representative of the application requirements, and which includes measurements points both near and far away from walls.

5.6.2 Latency and Energy Efficiency

Latency is a metric that represents the response time of the localization system, i.e. the time that system needs in order to update the location after the request for location estimation. Latency is measured by the time interval between the beginning and end of localization procedure of a node. Latency of the localization is an important metric because some localization use-cases such as emergency services require fast response time.

Energy efficiency is another metric which can be important particularly for wireless sensor networks (WSNs) where nodes must function completely wireless, and therefore are not connected to the power grid. The result of the measurement hardware is power and is expressed in Miliwatt (mW). The unit, defined as one joule per second, measures the rate of energy conversion or transfer ($W = \frac{J}{s}$). These measurements span a certain period but will only start once the system is up and running. For this metric, the same equations (average, min, max, ...) are derived as for the accuracy metric (Equations 5.3). Since the localization infrastructure is sometimes connected to power grid, a distinction is made between the power efficiency of the infrastructure and the clients.

5.6.3 Interference and Environmental Robustness

The RF-based indoor localization approaches are subject to exogenous interference caused by coexisting devices and technologies and endogenous interference caused by the other nodes using the same technology. The interference effect on the performance of localization schemes is measured by investigating the degradation of the accuracy under different interference circumstances. Different types and amounts of interference are specified and generated to study the interference robustness of this scheme, including competing wireless technologies but also microwaves and synthetic interference.

On the other hand, the RF indoor localization approaches are naturally exposed to the difficulties of indoor environment. Indoor environments are susceptible to changes caused by variation of network topology, room layout, walls, and channel conditions. The environment robustness determines if a solution is stable operating in different environments.

To measure these metrics, two different phases are necessary. The first phase is to calculate a metric from Section 5.6.2 in well-defined and controlled environments (e.g. accuracy with interference or latency with mobility). The second phase is to compare the mean, standard deviation, minimum and maximum value of this derived metric with the original performance metric.

5.6.4 Mobility, Scalability, Repeatability and Reproducibility

One important feature of wireless networks is the variable topology of the network due to the mobility and varying number of users in a given area. The mobility metric is defined as the variation of performance metrics with the speed of the localized node and characterizes how the performance of the localization schemes is changing from low-mobility regime to high-mobility one.

The scalability metric is concerned with the density of nodes and characterizes the performance of the localization schemes in sparse and dense networks. To measure these metrics, two scenarios are defined and compared with for each case corresponding respectively to low/high mobility regimes and low/high density regimes.

Repeatability implies that, if the same benchmark runs twice, results in the same score under well determined conditions. However, this equality is not strict in wireless benchmarking due to a certain level of indeterminism. For repeatability to apply, acceptable error margins should be formally defined. To evaluate this metric, the solution will be reinstalled multiple times in the same testbed under the same conditions and the variation in the accuracy is checked.

Reproducibility is an extension on repeatability, where, if the same benchmark runs twice on a different testbed or location that represents the same environment, it should produce the same results. The same error margins on equality apply as in repeatability.

5.7 Interference Robustness and Environmental Awareness

Environmental awareness and coexistence with other users and/or technologies is one of the core requirements of the Future Internet. This can be seen by the great push towards cognitive radio/dynamic spectrum access in the wireless research community, but also by the attempts to add awareness and cognitive features to existing standards. As such, a secondary goal of the EVARILOS project is the development and evaluation of localization solutions that add RF interference robustness to (existing) indoor localization solutions, such that indoor localizations also perform well in real-life Future Internet environments which are subject to uncontrolled interference.

One approach to enhance the robustness of indoor localization is to utilize the information gathered for environmental awareness and coexistence for the assessment of the quality of the localization process. A wireless device operating in the Future Internet will typically have detailed information about its spectral environment, either through spectrum sensing, or information retrieved from a database. Based on this information it can e.g. choose the best (i.e. least interfered) frequency to be used for the localization procedure. Alternatively, it can at least adjust the expected precision of the result based on the amount of expected interference.

The goal of the evaluation of interference robustness in EVARILOS project is adding a new class of approaches to RF-based localization to combat interference drawbacks. The solutions will be evaluated using the above described benchmarking methodology. We investigate to which extent such cognitive functionality of environmental awareness can improve the robustness of indoor localization against interference. From the investigations we will derive guidelines for the different classes of localization approaches on how to use which information to increase the interference robustness of indoor localization.

5.8 Conclusion

This paper presents an overview of the EVARILOS project which targets benchmarking and evaluation of indoor localization solutions. The general benchmarking methodology is presented. This methodology is a collection of scenarios that consist of environment and interference descriptions and different evaluation metrics. As the primary benchmarking metrics we define point and room accuracy, latency and energy efficiency. The secondary metrics, derived from them, evaluate the localization schemes under different environments, interference profiles, mobility, scalability and repeatability. By assigning different weight factors to these metrics, the benefits of different localization solutions for specific application domains can be compared objectively. During the project, selected localization solutions will be used as representative samples from different classes of existing RF-based indoor localization solutions.

The EVARILOS benchmarking methodology is currently being implemented on two testbeds belonging to the FP7 FIRE facility project CREW: Berlin testbed and Ghent testbed using IEEE 802.11, IEEE 802.15.1, and IEEE 802.15.4 technologies. We will experimentally apply the benchmarks to selected solutions on the two testbeds, in order to prove that the EVARILOS benchmarking methodology is generally applicable in different testbeds.

Once the EVARILOS benchmarking handbook is in a final version, a benchmarking suite will be developed in order to make an open call for participation possible. With this suite, an open call experimenter can test his localization solution in the two testbeds described above. In this way, a fair comparison between the experimenters' solutions can be made using the EVARILOS benchmarking handbook. The handbook includes a detailed list of metrics that determine the quality of the solution, together with well defined scenarios so that the experimenter has the all the information needed to perform the experiment and evaluate his solution in the testbeds provided by the EVARILOS project. Simultaneously there will start several tests of localization solutions with and without interference. These measurements will be used to fine tune the benchmarking handbook.

References

[1] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. *Range-free localization schemes for large scale sensor networks*. In Proceedings of the 9th annual international conference on Mobile computing and networking, pages 81–95. ACM, 2003.

[2] M. Keshtgary, M. Fasihy, and Z. Ronaghi. *Performance evaluation of hop-based range-free localization methods in wireless sensor networks*. ISRN Communications and Networking, 2011:9, 2011.

[3] S.-G. Zhang, Y.-P. Zeng, L.-J. Chen, D.-X. Chen, and L. Xie. *Performance evaluation of localization algorithms for mobile sensor networks*. Ruanjian Xuebao/Journal of Software, 22(7):1597–1611, 2011.

[4] N. B. Priyantha. *The cricket indoor location system*. PhD thesis, Massachusetts Institute of Technology, 2005.

[5] F. Winkler, E. Fischer, E. Graß, and G. Fischer. *A 60 GHz OFDM indoor localization system based on DTDOA*. In Proc. of the 14th IST Mobile and Wireless Communications Summit, pages 1–5. Citeseer, 2005.

[6] P. Bahl and V. N. Padmanabhan. *RADAR: An in-building RF-based user location and tracking system*. In INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, volume 2, pages 775–784. Ieee, 2000.

[7] M. Azizyan, I. Constandache, and R. Roy Choudhury. *SurroundSense: mobile phone localization via ambience fingerprinting*. In Proceedings of the 15th annual international conference on Mobile computing and networking, pages 261–272. ACM, 2009.

[8] D. Munoz, F. B. Lara, C. Vargas, and R. Enriquez-Caldera. *Position location techniques and applications*. Academic Press, 2009.

[9] X. An, J. Wang, R. V. Prasad, and I. Niemegeers. *OPT: online person tracking system for context-awareness in wireless personal network*. In Proceedings of the 2nd international workshop on Multi-hop ad hoc networks: from theory to reality, pages 47–54. ACM, 2006.

[10] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *Automated linear regression tools improve rssi wsn localization in multipath indoor environment*. EURASIP Journal on Wireless Communications and Networking, 2011(1):1–27, 2011.

[11] Y. Kwon, K. Mechitov, S. Sundresh, W. Kim, and G. Agha. *Resilient localization for sensor networks in outdoor environments*. In Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on, pages 643–652. IEEE, 2005.

[12] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'dea. *Relative location estimation in wireless sensor networks*. Signal Processing, IEEE Transactions on, 51(8):2137–2148, 2003.

[13] F. Vanheel, J. Verhaevert, E. Laermans, I. Moerman, and P. Demeester. *A linear regression based cost function for WSN localization*. In Software, Telecommunications and Computer Networks (SoftCOM), 2011 19th International Conference on, pages 1–5. IEEE, 2011.

[14] N. J. Gordon, D. J. Salmond, and A. F. Smith. *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*. In Radar and Signal Processing, IEE Proceedings F, volume 140, pages 107–113. IET, 1993.

[15] R. E. Kalman. *A new approach to linear filtering and prediction problems*. Journal of basic Engineering, 82(1):35–45, 1960.

[16] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson. *Easytracker: automatic transit tracking, mapping, and arrival time prediction using smartphones*. In Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, pages 68–81. ACM, 2011.

[17] C. Morelli, M. Nicoli, V. Rampa, and U. Spagnolini. *Hidden Markov models for radio localization in mixed LOS/NLOS conditions*. Signal Processing, IEEE Transactions on, 55(4):1525–1542, 2007.

[18] D. Madigan, E. Einahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, and A. Krishnakumar. *Bayesian indoor positioning systems*. In INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, volume 2, pages 1217–1227. IEEE, 2005.

6

Platform for Benchmarking of RF-based Indoor Localization Solutions

The previous chapter described a benchmarking methodology which (partially) solves the issues identified in Chapter 3 and 4 of Part II. However in order to address these issues more efficiently, benchmarking tools aligned with the methodology (described in Chapter 5) are desirable. Therefore, this chapter describes a web-based benchmarking platform containing multiple tools which can be used by researchers to validate their solution in an objective, proper way. The final goal of this platform is to create a community of researchers wherein they can evaluate and compare their solution in a more uniform way. This benchmarking platform is available at http:// evarilos.intec.ugent.be.

⋆ ⋆ ⋆

T. Van Haute, E. De Poorter, F. Lemic, V. Handziski, N. Wirström, A. Wolisz, I. Moerman.

Published in IEEE Communications Magazine, September 2015.
[IF: 4.007 , Q1]

Abstract Over the last years, the number of indoor localization solutions has grown exponentially, many of which use a wide variety of different technologies and approaches. Unfortunately, there is currently no standardized evaluation method for comparing their performance. As a result each solution provider evaluates their solution in their own, proprietary environment using proprietary evaluation metrics. Consequently, it currently is an impossible task to compare the performance of multiple localization solutions with each other. To remedy this situation, this paper presents the EVARILOS Benchmarking Suite (EBS), a platform whereby multiple solutions (based on different technologies) can automatically be compared and evaluated in different environments and using multiple evaluation metrics. To this end, a testbed independent benchmarking suite was created, in combination with multiple tools for creating, storing and comparing performance results of multiple indoor localization solutions. This platform implements the standardized evaluation methods described in the EVARILOS Benchmarking Handbook (EBH), which is aligned with the upcoming ISO/IEC 18305 standard: "Test and Evaluation of Localization and Tracking Systems". The suite and tools can be used real-time on existing wireless testbed facilities, but also offer a remote offline evaluation method using pre-collected data traces. Finally, by analysing and comparing the performance of three different indoor localization solutions, the paper demonstrates the need for such an objective evaluation and testing method that considers multiple evaluation criteria in different environments.

6.1 Introduction

This paper addresses one of the major problems of indoor localization research: the lack of comparability between existing localization solutions due to the fact that numerous published solutions have been evaluated under individual, not comparable and not repeatable conditions. These facts are partly caused by the complexity required for correctly evaluating an indoor localization solution, which requires technical expertise to efficiently setup large-scale experiments, to control the experimental environment, to gather the necessary performance data and to calculate output metrics using standardized methods.

Going through these steps is time consuming, and more theoretically inclined researchers typically lack the necessary technical skills to perform these steps efficiently and accurately. This paper addresses these deficiencies by providing solutions that allow easy set-up and evaluation of indoor localization experiments. The main contributions described in this paper are as follows.

- This paper describes a generic benchmarking suite that implements the standardized evaluation methods described in the EBH, and which is aligned with the upcoming International Organization for Standardization / International Electrotechnical Commission (ISO/IEC) 18305 standard “Test and Evaluation of Localization and Tracking Systems”.

- Scripts are available for instantiating the components of the benchmarking suite on multiple Future Internet Research and Experimentation (FIRE) wireless test facilities [1].
- Finally, to simplify the act of benchmarking and evaluating a localization solutions, open datasets are provided containing environmental Radio-Frequency (RF) data from multiple environments that can be used as input for the benchmark suite components, thereby allowing offline evaluation of localization solutions without requiring any technical knowledge about configuring and utilizing wireless test facilities.

This paper is structured as follows. Section 6.2 contains the related work. In Section 6.3 the benchmarking suite is further explained in detail. The integration of the EBS in a wireless test facility and the public datasets can be found, respectively, in Section 6.4 and 6.5. Section 6.6 demonstrates the benchmarking suite with an experimental validation. Finally, Section 6.7 contains a conclusion.

6.2 Related Work

Since the number of indoor localization solutions is growing, also the awareness that a more thorough way of evaluating and comparing is necessary. A well defined objective evaluation methodology is in need of a wide range of metrics. Some metrics are important for a theoretical point of view and as such are well-suited for analyzing and improving algorithms of researchers, whereas other focus on the performance of end-systems and as such are more important for the industry. If only accuracy is taken into account, then the results can give a distorted view. Therefore, Hui Liu et al. states in [1] that precision, complexity, scalability, robustness and cost should be included if comprehensive performance analysis is required.

Additionally, [1] also recognizes the lack of a decent and objective evaluation methodology for localization solutions. Towards this end, the following organizations are trying to develop more standardized testing approaches for localization solutions.

- First, the FP7 EVARILOS project focuses on the **EVA**luation of **R**F-based **I**ndoor **LO**calization **S**olutions and published the “EVARILOS Benchmarking Handbook” [2]. The handbook describes which evaluation metrics are important to consider during the evaluation phase

[1] http://www.ict-fire.eu/home.html

(such as point accuracy, response time, energy consumption, etc.) and how to calculate them. Furthermore, this book contains a set of scenarios which describe how to adequately evaluate an indoor localization solution. The project is also the first to point out that current scientific literature lacks studies on the effect of interference on indoor localization solutions, although interference is expected to be present at most sites where these systems are installed. The EBH includes a wide range of evaluation metrics, including functional metrics such as response delays and deployment metrics such as setup time and required infrastructure.

- Recently, the ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) have established a joint technical committee, ISO/IEC JTC 1 so they can work together on the ISO/IEC 18305: "Test and Evaluation of Localization and Tracking Systems" [2]. Current drafts include evaluation methodologies for a single technology (e.g. Bluetooth), as well as methodologies for evaluation full localization systems, which is more in line with the EVARILOS project. In contrast to the EVARILOS project, this work also includes a wide range of other non-RF based technologies such as motion sensors, but does not yet include non-accuracy related metrics such as ease-of-use or energy consumption. At the time of writing, none of the drafts are publicly available yet.

- Finally, in the EvAAL project (Evaluating AAL Systems through Competitive Benchmarking) [3], a competition is held that aims at establishing benchmarks for comparing Ambient Assistant Living (AAL) solutions. For this competition, besides accuracy, also usability metrics are defined such as installation complexity, user acceptance, availability and interoperability with AAL Systems. In contrast to the EVARILOS project, the evaluation process is not automated, and requires bringing physical devices to a target site.

Most scientific papers evaluate their solution in an easily accessible building near to the development area of the authors. Typically, this is an office environment with brick walls [3, 4]. Since the evaluation is rather time consuming, most localization solutions are evaluated only in a single environment. Both the EVARILOS project and ISO/IEC JTC 1 refer to the fact that this is not representative for many buildings. As such, our platform offers the user the possibility to download datasets recorded in multiple environments: an office environment with brick walls, an office environment with plywood walls and finally an industrial-like open environment. In addition, since the accuracy strongly depends on the used evaluation points,

[2] http://www.iso.org
[3] http://evaal.aaloa.org

e.g. points near a wall versus open spaces, our public datasets that are available contains data measured at a wide range of measurement points.

6.3 Benchmarking Platform

This section describes the EVARILOS Benchmarking Suite. The EBS has been created to cope with the fact that, although a significant number of experimental testbed facilities is available [5, 6], evaluating the performance of a localization solution under controlled conditions using standardized performance metrics has proven to be very complicated for researchers that have no -or limited- experience with experimental research. The EBS solves this issue by providing an open software solution that implements user friendly methods to realize the full performance analysis cycle.

The benchmarking suite described in this section implements the standardized evaluation methods described in the EBH and is aligned with the upcoming ISO/IEC 18305 standard “Test and Evaluation of Localization and Tracking Systems”. The developed software components are independent of any experimental facilities and use open data principles, allowing researchers to download and modify any of the components.

An overview of the EBS architecture is shown in Figure 6.1.

- The *rectangles* represent components that are available as web services. These components run on a cloud service where they can be accessed, or they can be downloaded to modify and/or run locally.
- The *parallelograms* represent data structures that are used to exchange data between the web services.
- Finally, the *flags* represent the tools that can be used to analyze and visualize the different steps of the process.

The architecture consists of a set of components that, when used sequentially, implement a workflow that represents three experimentation steps. A summary can be found below, in the next subsections, each step will be discussed more in detail:

1. During a pre-experimentation phase, users can download environment-specific training datasets from the public repositories. These datasets are typically used for training the localization solution.
2. Next, the experiments are orchestrated. Tools are available for the automatic generation of experiment configurations, including specifications of the used evaluation points, the interference patterns that will be generated, etc. Based on these descriptions, experiment

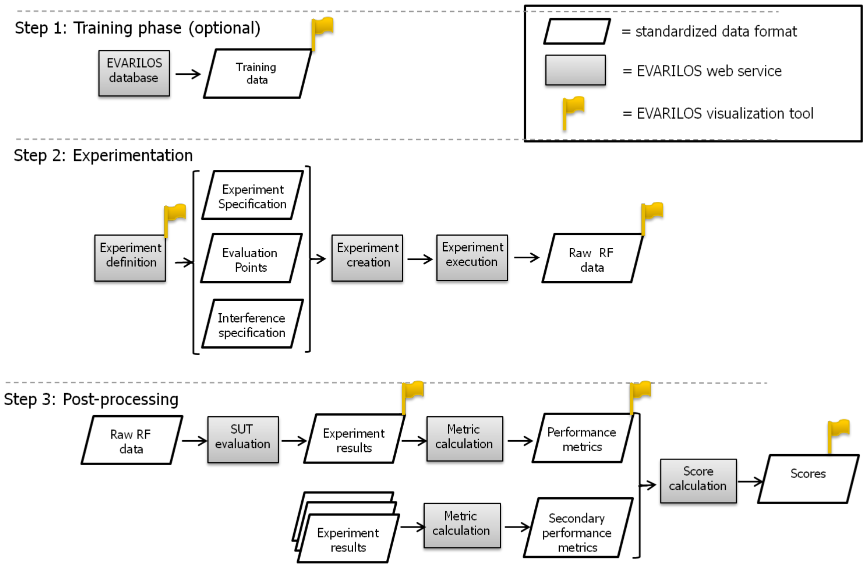


Figure 6.1: Overview of the components of the EBS and the data structures used to exchange information between the components

scripts are created (for example with cOntrol and Management Framework (OMF) [4] scripts which are used in many recent wireless testbeds [5]) and are automatically executed. Note that this step can be omitted if the next step utilizes pre-collected input (e.g. Wi-Fi beacons) for the localization solution.

3. Finally, the environmental RF data is fed to the System Under Test (SUT), either real-time or using pre-recorded measurements depending on the experiment configuration. The estimated locations are stored together with additional performance metrics such as the response delay. It is also possible to combine results from multiple experiments to observe how certain evaluation metrics evolve over multiple experiments.

6.3.1 Training Phase

The training phase offers experimenters the possibility to train their localization solutions based on measurements that are performed in advance in a representative location. Measurements for training purposes are captured in an area that is representative for the experimentation phase, typi-

[4] http://omf.mytestbed.net/projects/omf6/wiki/Wik
[5] http://mytestbed.net/projects/omf/wiki/DeploymentSite

cally the data is captured in the same environment where the SUT will be evaluated.

However, the data does not necessarily correspond to the data that will be used during the evaluation. It is possible that the data is (i) captured at a different time and/or (ii) captured using devices from a different manufacturer and/or (iii) is captured at different evaluation points than the one used during the performance evaluation. It is important to differ in at least one of these aspects, otherwise the SUT is trained to be biased towards the performance evaluation of step 2 of the process.

The suite offers researchers a database to access previously measured environmental information relevant for their localization solution. Users can either download the data directly from the EVARILOS data repository or can access an EVARILOS API that encapsulates the data and can serve the data in a finer granularity.

6.3.2 Experimentation Phase

The experimentation phase offers experimenters the possibility to define setups for evaluating experiments, as well as an interface for automatic execution of data collection or localization experiments in FIRE facilities. The user will start with an “experiment definition” (see Figure 6.1). The role of the experiment definition component is to configure all aspects of the experiment that will be used to evaluate a SUT.

To this end, the experiment definition components requires the following input: the *experiment specification* (e.g. which nodes will be used as anchor points, when will the experiment be scheduled, which binary files to use, etc.), the *evaluation points* (at which locations is the SUT evaluated) and the *type of (artificial) interference* that should be generated. To assist with this process, a fully automated web service is available where users can select amongst different preconfigured options. Of course, it is possible to modify any of the default settings to adjust the experiment behavior. This information is also stored in standardized data formats.

Next, the “experiment creation” component is executed, this is a fully automated step whereby the testbed independent information will be translated into testbed dependent scripts (at the moment, only OMF for w-iLab.t II in Zwijnaarde). The final step is the actual execution of the experiment. In this step, the translated scripts are executed on the corresponding testbed. The outcome of this execution is stored in a standardized data format as well.

6.3.3 Post-processing Phase

During the final phase of the EBS, the location estimates are processed and performance metrics are calculated. The environmental RF data, resulting from the experiments during the previous step, are fed to the SUT. During this step, either real-time input data is used, coming from the experimentation step, or previously collected data is replayed.

The input data includes data such as Wi-Fi beacon information, IEEE 802.15.4 beacons, etc. and is used by the SUT to generate position estimates. The provided location estimates are stored in the experiment results data structure, together with additional metadata.

Experiment results from multiple experiments can be combined to observe how certain evaluation metrics evolve over multiple experiments. These results are stored in the secondary metrics data structure. For comparability purposes, a final score can be assigned to a SUT performance. This score is an abstraction of the performance of a SUT in a specific environment and necessarily hides many intrinsic trade-offs.

Finally, it is worth noting that the full post-processing phase can also be applied to location estimates from non-EBS compliant solutions. As long as the experiment results are provided in the correct data format, the tools can be used to analyse and rank the outcome of any localization solution.

6.4 Integration of the EBS in a Wireless Experimentation Facility

The EVARILOS Benchmarking Suite is designed to simplify the evaluation of RF-based localization solutions, also referred to as SUTs. The tools can be used “as-is” by utilizing pre-collected data traces as input. However, the benchmark components can also be used to *facilitate the evaluation of localization solutions in new environments*. The available deployment options for indoor localization benchmarking are presented in Figure 6.2. Three main components can be identified.

- The bottom layer represents a wireless experimentation facility or testbed.
- The components of the EBS suite are installed on a server in the test facility. The EBS includes software tools that facilitate wireless experimentation and evaluation of localization solutions (see Section 6.3).
- Finally, the upper layer represents the SUT, which can include both hardware and/or software components.

The EBS is also integrated in existing FIRE facilities. This integration is part of the “experiment execution” component illustrated in Figure 6.1. Automatic conversion from experiment descriptions to OMF scripts is supported, thereby integrating and simplifying the complex steps that otherwise needed to be taken for objective experimentation. Building on top of the CREW Cognitive Radio testbeds [6], the infrastructure leverages a robotic mobility platform, which serves as a reference localization system and can transport the localized device in an autonomous and repeatable manner.

In addition, the suite uses the capabilities of the CREW testbed infrastructure to generate typical interference scenarios in a replicable manner. This further improves benchmarking of an indoor localization system by testing SUT performance under realistic and repeatable interference conditions.

Interaction between the SUT and the EBS is designed to be as simple as possible: only two REST [7, 8] interfaces are necessary. One to provide location estimates to the EBS and an optional one to provide feedback data to the SUT.

- The main responsibility of the web interface is to request location estimates. From time to time, the EBS requests the location from the SUT through a REST web service. As such, the minimum requirement for the SUT to use the EBS is to provide the location estimation over HTTP when requested.
- Optionally, the EBS can provide the SUT with prerecorded or real-time environment data (such as RSSI values, time of arrival information, etc.) through a second REST web service. This data can be collected and can at a later time be offered to future experimenters as an open data set.

This architecture allows experimenters to choose amongst different utilization options.

Option 1: the evaluation of a localization algorithm using real-time or prerecorded data from a testbed. In this scenario, the localization algorithms can run remotely from the experimentation facility.

Option 2: the evaluation of a localization solution using software running on an existing wireless testbed. In this scenario, the localization algorithms can run on local hardware that is available at the experimentation facility.

Option 3: the evaluation of localization hardware using a testbed. In this scenario, experimenters can install custom hardware at the experimentation facility whilst still using the EBS components for evaluation their solution.

One of the major advantages of the EBS is that all three approaches make use of the same common components. The feasibility of all three options has been demonstrated through the EVARILOS Open Challenge [9]

[6] http://www.crew-project.eu/

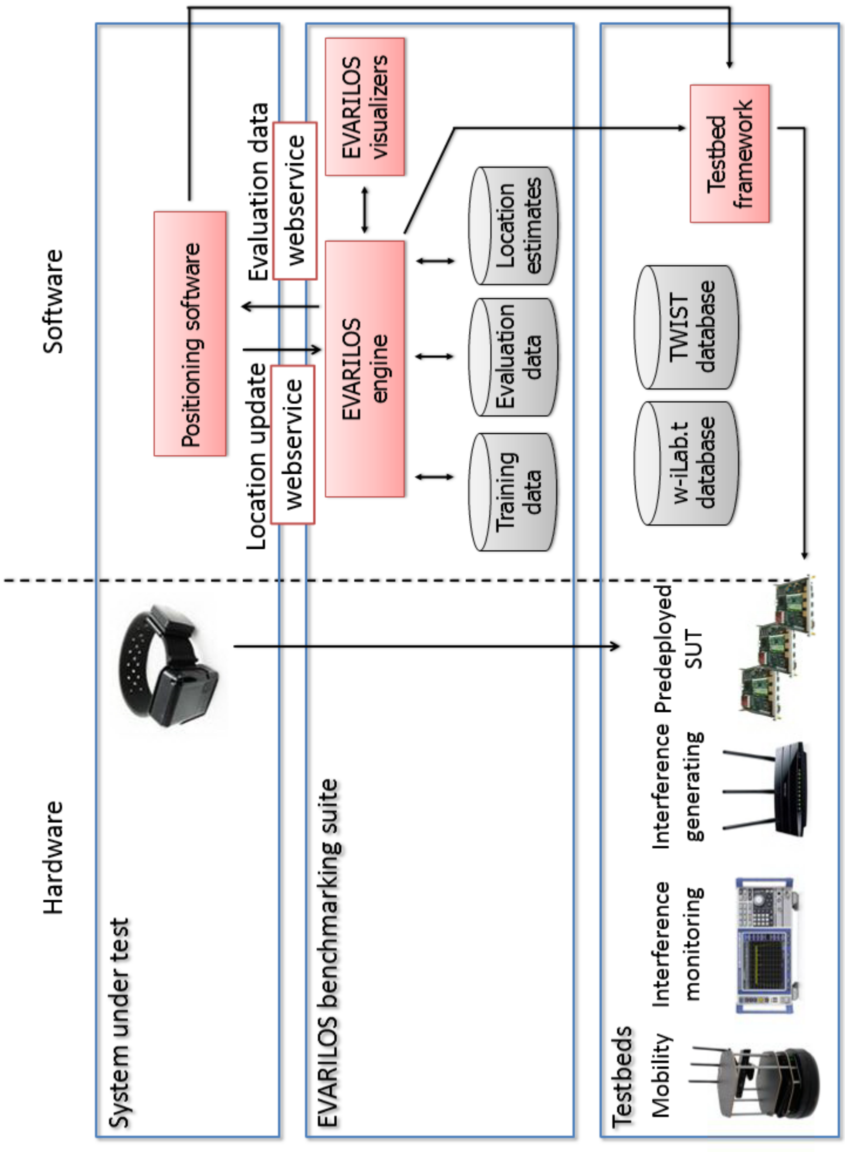


Figure 6.2: Deployment of the EVARILOS benchmarking suite

as well as during the Microsoft Indoor Localization Competition (IPSN 2014) [10].

6.5 Public Datasets

One of the features of the EBS is the capability to replay previously recorded RF-data for offline evaluation of localization solutions. This feature solves one of the main challenges for the indoor localization research community, namely the complex and expensive process of obtaining relevant RF-features from multiple environments. EVARILOS offers a wide range of available pre-recorded RF-data sources through its user interface. However, for those researchers that prefer to download full annotated datasets, EVARILOS also offers the possibility to retrieve separate datasets for research purposes. Two types of datasets are especially relevant: RF data and performance information.

6.5.1 RF Data

Environmental RF-data can be used as a basis either for training an algorithm (e.g. by creating propagation models) or for offline evaluation of a SUT. EVARILOS makes available the measured RF data from multiple environments, including a plywood office environment (w-iLab.t I [5]), a brick office environment (TWIST [11]), an industrial-like environment (w-iLab.t II [5]), a hospital environment and an underground mine.

To evaluate a solution in a wide range of conditions, the datasets contain significantly more data than would be used in a typical operational environment. The datasets are rich in terms of number of collected samples per evaluation point (over a thousand samples per evaluation point), the captured data types (including Wi-Fi beacons, sensor RSSI and sensor time-of-flight information), the used configuration settings (multiple frequencies, multiple transmission powers) and the used anchor points (data is collected from up to 60 anchor points per evaluation point). This richness of the dataset makes the data relevant for a wide range of interested researchers and allows investigating how changing any of these parameters influences the performance of the solution. Transforming the overdimensioned dataset into a set that is more sparse (and more realistic from an operational point of view) can easily be done by removing any unnecessary information (subsampling). In addition, the available environment data is annotated with metadata describing the exact conditions in which the data was captured. This metadata describes characteristics such as the used hardware, the height of the antenna, the type and angle of the antenna, timestamps, measurement frequency, etc.

6.5.2 Performance Information

EVARILOS gives a ranked overview of evaluated solutions on its webpage. However, these performance indicators necessarily hide a number of low-level statistics. Researchers interested in evaluating also the temporal behavior or spatial behavior of EVARILOS solutions can analyze the performance datasets. EVARILOS makes available the results from its own localization solutions, as well as of those solutions that participated in the open challenge. Each of these datasets also has its associated experiment configuration settings, allowing detailed analysis of not only the performance but also the conditions in which the solutions were evaluated.

6.6 Experimental Validation

Finally, the need for a standardized evaluation method will be demonstrated by showing that the performance of localization solutions depends strongly on the environment of the evaluation facility and can only be done by including multiple evaluation metrics.

6.6.1 Three Indoor Localization Systems

In order to develop, test and optimize our suite, three different indoor localizations solutions were used as SUTs. The first solution is designed by N. Wirström et al. [12]. The basic concept behind this localization solution is the following: measurements are performed by letting a stationary node transmit packets to the testbed nodes that reply with a hardware ACK (acknowledgment). The initiating node measures both the time between the transmission of the packet and the reception of the ACK, and stores the RSSI values associated with the ACK. These measurements are then processed using *Spray*, a particle filter based platform. The basic idea of the ToF ranging is to estimate the distance between two nodes by measuring the propagation time that is linearly correlated to the distance between the nodes when they are in LoS.

A second solution is provided by F. Lemic [13] which is based on fingerprinting. Fingerprinting methods in the indoor localization are generally divided in two phases. The first phase is called the training or offline phase. In this phase, the localization area is divided in a certain number of cells. Each cell is scanned a certain number of times for different signal properties, and using a methodology for processing the received data a fingerprint of each cell is created. By using the obtained training fingerprints the training database is created and stored on the localization server. In the second phase, known as the runtime or online phase, a number of scans of the environment are created using the user's device. From the scanned data, using the same predefined data processing methodology, the runtime fingerprint is created and sent to the localization server. At the server's side the runtime fingerprint is compared with the training dataset using the matching method. The training fingerprint with the most similarities to the runtime fingerprint is reported as the estimated position.

A final localization solution [14] that has been implemented and evaluated is a hybrid combination of a range-based and a range-free algorithm. It includes a range-based location estimator based on weighted RSSI values. Each RSSI value can be matched with a certain distance. The proposed algorithm in [14] not only uses the RSSI values to measure the distance between a fixed and mobile node, but also the distance between the fixed nodes. These values function as weight factors for the distance calculation between the fixed and mobile node. Once the distances are known, triangulation can be applied in order to determine the final position of the person / object that needs to be localized. This approach is combined with a range-free algorithm, this does not take RSSI-values into account. If a mobile sensor node has a range of 10 meters, then a fixed node can only receive his messages if the mobile node is maximum 10 meters away. This is the only information that is used to calculate the position of a mobile node. For this approach, it is important that the transmission power is well configured. If the power is too low, the mobile node could be out of range between two fixed nodes. On the other hand, if the power is too high, too many fixed nodes will receive the beacon and a wrong estimation could be made.

6.6.2 Analysis of a Single Solution

A first major advantage of the EVARILOS benchmarking suite is its capability to streamline the process for giving a researcher a better insight in his solution. Every solution contains a set of adjustable parameters which can considerably influence the overall performance of the solution, optimizing this set of parameters can be a hard task for the researcher. Therefore, the EVARILOS benchmarking suite can easily compare the same solution, using multiple values for one single parameter.

This can be demonstrated with a small example. The hybrid solution [14] described in the section above states that the transmission power is an important value that needs to be configured well in order to receive acceptable results. Therefore, the solution was evaluated using the EBS using multiple transmission powers, the outcome of which is shown using a Cummulative Distribution Function (CDF) (Figure 6.3) and a table with multiple metrics (Table 6.1). Based on these results, it is clear that this solution obtains the lowest average error when the transmission power equals three. But it also illustrates inherent trade-offs that are present in the solution: suppose the response time would be the most important criteria, then a transmission power of 31 would be the best option. This examples illus-

Table 6.1: Statistical information about the performance of the hybrid algorithm in TWIST testbed

Metric	Tx3	Tx7	Tx19	Tx31
Average error [m]	4.63	7.08	6.93	8.31
Min. error [m]	0.75	0.83	0.80	0.82
Max. error [m]	10.20	17.52	18.93	19.31
Median error [m]	4.39	6.81	6.68	8.63
RMS error [m]	5.13	7.75	7.82	9.24
Room accuracy [%]	26.67	6.70	13.45	9.56
Response time [ms]	1503	1507	480	460

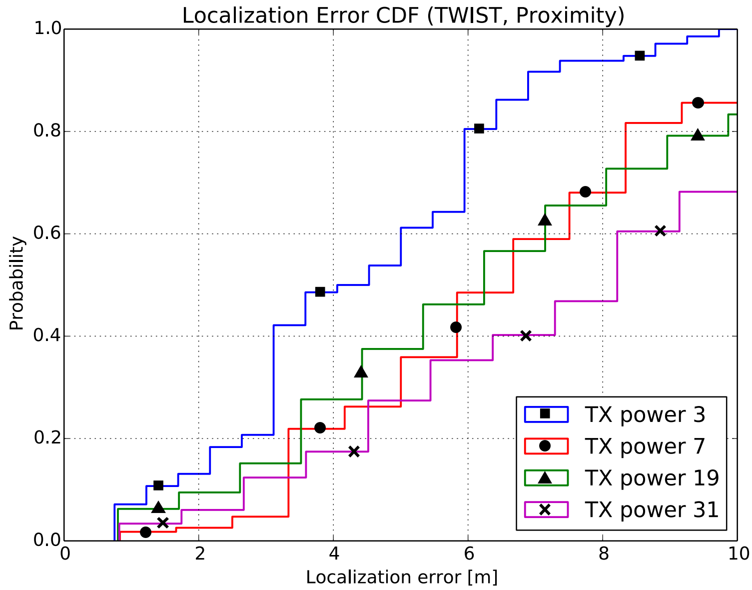


Figure 6.3: CDF of the proximity based solution in TWIST

trates the advantages of the EBS for fast and efficient identification of an optimal operating point depending on adjustable parameters, and demonstrates the need for considering multiple metrics to identify trade-offs.

Table 6.2: TWIST (Berlin) Summarized benchmarking results

Algorithm	Average Error [m]	Room accuracy [%]	Response time [ms]	Energy efficiency [mW] Mobile	Energy efficiency [mW] Fixed
Particle filter solution					
Spray RSSI	4.35	45.00	14 285	∼ 105	∼ 105
Spray ToA	5.56	30.00	14 282	∼ 105	∼ 105
Fingerprinting solution					
KL Distance	2.7	50.0	∼ 35 000	∼ 7000	∼ 500
ED Distance	2.2	80.0	∼ 35 000	∼ 7000	∼ 500
PH Distance	2.0	85.0	∼ 35 000	∼ 7000	∼ 500
Hybrid solution					
TX Power = 3	4.6	26.7	1 503.1	∼ 30.9	∼ 47.4
TX Power = 7	7.1	6.7	1 507.6	∼ 35.1	∼ 47.4
TX Power = 19	7.9	13.4	480.6	∼ 47.1	∼ 47.4
TX Power = 31	8.7	9.5	460.9	∼ 57.6	∼ 47.4

6.6.3 Comparison Between Multiple Solutions

Finally, Table 6.2 compares the performance of three different solutions evaluated using the EBS by considering multiple evaluation criteria. By utilizing the same evaluation points, objective comparisons are possible. Again, the results illustrate the presence of trade-offs that can only be observed by objectively comparing multiple metrics.

More specifically, it demonstrates that the approach taken in most current scientific papers, wherein point accuracy is considered the main metric, fails to take into account that the most accurate solution in this case is also the one that consumes most energy and has a significantly higher response time than the other solutions.

6.7 Conclusion

The proliferation of RF-based indoor localization solutions raises the need for testing systems that enable objective evaluation of their functional and non-functional properties. Although a significant number of localization solutions is available, the evaluation of these solution use different approaches in terms of used performance metrics and evaluation methodology.

This paper tries to solve these shortcomings by providing tools for evaluating and comparing localization solutions using the standardized evaluation methods described in the EBH, and which is aligned with the upcoming ISO/IEC 18305 standard “Test and Evaluation of Localization and Tracking Systems”.

The paper describes three main contributions: (i) a benchmarking suite, (ii) scripts to integrate this suite in FIRE experimental facilities, and (iii) an open dataset to bootstrap efficient experimentation design and comparison of solutions.

This paper introduced a testbed independent benchmarking suite for automatized benchmarking of RF-based indoor localization solutions. Using a well-defined interface, the infrastructure obtains location estimates from the SUT, which are subsequently processed in a dedicated metrics computation engine.

The components can be accessed through web services that are available for external users or can be downloaded for custom modifications. The benchmarking suite has shown to be useful for locations where no testbed facilities are available. Multiple components of the suite were extensively used during the Microsoft Indoor Localization Competition (IPSN 2014) as well as the EVARILOS Open Challenge. In these events, the components of the benchmarking suite improve the time efficiency and ease of use of the experiments, as well as resulted in more objective comparability.

Finally, to accommodate the need for a wider accessibility of experimental data, open datasets are provided. These datasets include both annotated localization data from multiple environments, as well as detailed descriptions of the setup and outcome of the performed localization experiments from earlier experiments.

These repositories can be used to quickly evaluate a SUT in different environments, to analyse the effects of changing configuration settings, to analyse the setup of different experiments and to compare the performance of a wide range of localization solutions.

References

[1] H. Liu, H. Darabi, P. Banerjee, and J. Liu. *Survey of Wireless Indoor Positioning Techniques and Systems*. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37(6):1067–1080, 2007.

[2] T. V. Haute et al. *The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions*. In MERMAT 2013, 5 2013.

[3] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan. *Indoor Localization Without the Pain*. In Proceedings of the sixteenth annual international conference on Mobile computing and networking, pages 173–184. ACM, 2010.

[4] E. Martin, O. Vinyals, G. Friedland, and R. Bajcsy. *Precise Indoor Localization using Smartphones*. In Proceedings of the international conference on Multimedia, pages 787–790. ACM, 2010.

[5] S. Bouckaert, W. Vandenberghe, B. Jooris, I. Moerman, and P. Demeester. *The w-iLab. t Testbed*. In Testbeds and Research Infrastructures. Development of Networks and Communities, pages 145–154. Springer, 2011.

[6] F. Lemic, J. Büsch, M. Chwalisz, V. Handziski, and A. Wolisz. *Infrastructure for Benchmarking RF-based Indoor Localization under Controlled Interference*. In UPINLBS'14, Corpus Christi, Texas, USA, 2014.

[7] F. Lemic and V. Handziski. *Data Management Services for Evaluation of RF-based Indoor Localization*. Technical Report TKN-14-002, Telecommunication Networks Group, 2014.

[8] F. Lemic. *Service for Calculation of Performance Metrics of Indoor Localization Benchmarking Experiments*. Technical Report TKN-14-003, Telecommunication Networks Group, 2014.

[9] F. Lemic, V. Handziski, A. Wolisz, T. Constambeys, C. Laoudias, D. Zeinalipour-Yazti, C. Panayiotou, S. Adler, S. Schmitt, H. Will, T. Hillebrandt, M. Kyas, Y. Yang, and Y. Zhao. *EVARILOS Open Challenge: Track 3*. Technical Report TKN-14-005, Telecommunication Networks Group, Technische Universität Berlin, 2014.

[10] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, S. Sen, and V. Handziski. *Microsoft Indoor Localization Competition: Experiences and Lessons Learned*. In SIGMOBILE Mobile Computation and Communication Review (MC2R). SIGMOBILE, 2014. To appear.

[11] V. Handziski et al. *TWIST: A Scalable and Reconfigurable Testbed for Wireless Indoor Experiments with Sensor Network*. In RealMAN 2006, 2006.

[12] N. Wirström, P. Misra, and T. Voigt. *Spray, Embracing Multimodality*. 2014.

[13] F. Lemic. *Benchmarking of Quantile based Indoor Fingerprinting Algorithm*. Technical Report TKN-14-001, Telecommunication Networks Group, Technische Universität Berlin, 2014.

[14] T. Van Haute, J. Rossey, P. Becue, E. De Poorter, I. Moerman, and P. Demeester. *A Hybrid Indoor Localization Solution Using a Generic Architectural Framework for Sparse Distributed Wireless Sensor Networks*. In Computer Science and Information Systems (FedCSIS'14), pages 1009–1015. IEEE, 2014.

7

Benchmarking of Localization Solutions: Guidelines for the Selection of Evaluation Points

Finally, this chapter deals with the selection of evaluation points. It tackles one of the issues described in Chapter 3: the influence of the amount of used evaluation points. Therefore, guidelines are elaborated which try to formulate an answer to the following question: "How many evaluation points are required in order to represent an objective mean value of the error distances?". The outcome of this work enriches the benchmarking methodology described in Chapter 5.

⋆ ⋆ ⋆

T. Van Haute, E. De Poorter, E. Laermans, I. Moerman.

Submitted to Elsevier Ad Hoc Networks, May 2016. [IF: 1.66 , Q2]

Abstract Indoor localization solutions are key enablers for the realization of new services in multiple application domains. As such, a large body of scientific indoor localization research is currently being published. However, an objective evaluation methodology to report on the accuracy of these solutions is currently lacking. Researchers typically use only a limited number of evaluation points which are frequently selected based on convenience of access. There is currently no information on how this evaluation point

selection procedure impacts the reported accuracy. To remedy this, this paper evaluates the influence of using different number of evaluation points for the evaluation of localization solutions. Based on simple statistical parameters, the paper defines an estimator for calculating the confidence interval of estimated accuracy of evaluated localization solutions based on the number of used evaluation points. The obtained estimators are experimentally validated for multiple indoor localization solutions using different technologies (IEEE 802.15.4 and IEEE 802.11 Wi-Fi), for different localization algorithms and for multiple deployments (including an industrial environment and a hospital environment). The outcomes of this paper allow researchers to not only report the mean accuracy of evaluated localization solutions, but also to calculate the confidence with which these reported results are correct. This information can also be used as a stopping criteria during the evaluation of localization solutions, helping evaluators to decide if more evaluation points are required to obtain reliable results.

7.1 Introduction

In recent years, location-based information becomes indispensable in multiple application domains. As such, the number of published research focusing on indoor localization has increased tremendously. The outcome of these publications are mostly promising with ever-increasing accuracy results. However, the reported performance results are often biased since a standardized evaluation procedure is not applied. Each solution is evaluated in a different environment using custom-defined evaluation methodologies. However, the chosen environment and evaluation points can drastically influence the reported accuracy of the evaluated solution. As a result, it is hard to assess the reliability of the reported performance.

Within this paper, ‘an evaluation point’ is defined as *‘a physical location in a test environment whereof the exact coordinates are known, and that is used as a reference or baseline point’*. By comparing the coordinates of these evaluation points with the corresponding estimated positions from the localization solution, an indicator of the typical error is obtained. For practical evaluation purposes, typically measurement data is collected. The localization algorithm has to process this data and try to estimate the coordinates of this physical location, without having knowledge of the ground truth nor history-based information. The euclidean distance between the estimated point and the evaluation point is defined as the ‘error distance’.

To allow better insight in the impact of evaluation point selection on the reported accuracy, this paper aims to answer the following questions. (i) Does the number of evaluation points have an impact on the reported accuracy? (ii) If so, is their an way to determine how many evaluation points are required so that a reliable accuracy estimation can be obtained? (iii) Are there guidelines which researchers can follow in order to improve the reliability of their presented performance results. (iv) Finally, verify the correctness of the guidelines using multiple configurations.

The remainder of this paper is structured as follows. Section 7.2 contains the related work. Section 7.3 describes in detail our used testbeds the evaluated indoor localization algorithms. The behaviour of the typical error distances is analysed in Section 7.4. The next section, Section 7.5 investigates the behaviour and properties of the mean value and standard deviation of the error distances. Based on these findings, guidelines regarding to the optimal number of evaluation points and the confidence interval are provided in Section 7.6. Finally, the paper's conclusions can be found in Section 7.7.

7.2 Related work

There exists a large number of papers describing and evaluating indoor localization solutions. However, many solution's performance analysis is performed using insufficient evaluation points resulting in biased results. This section discusses the current state of the art approaches for evaluating indoor localization solutions, both in scientific literature and during open indoor localization competitions.

7.2.1 Evaluation procedures in scientific literature

In [1], the authors analyse and compare 21 research papers each presenting an Indoor Localization Sensing (ILS) solution. The paper summarizes the algorithm design, devices, test setup and the performance results of each of these. However, the number and locations of the evaluation points are not provided. As such, it becomes hard to objectively compare the proposed solutions.

In contrast, Xiaowei Luo et al. present in [2] a comparative evaluation of Received Signal Strength Indicator (RSSI) based indoor localization techniques for construction job sites. Four different algorithms (MinMax, Maximum Likelihood, Ring Overlapping Circle RSSI & k-nearest Neighbour) are each evaluated in two different test environments (Building & Job site). The authors selected 21 measurement points in the first testbed (44.87 m^2) and 18 points in the second testbed (32.26 m^2). Although the number of evaluation points is provided it is unclear where the measurement points were located in the test environment, and if this number was sufficient to reliably obtain mean accuracy results.

As an alternative, Gayathri Chandrasekaran et al. present an empirical evaluation of the limits on localization using signal strength [3]. In their work, a trace-driven emulation is used to evaluate the performance of 12 different localization algorithms. They applied the “leave-one-out” approach to evaluate the algorithms using a fingerprint database. As such, a dataset of 400 points is split into 400 sets of 399 training points and one testing point. They presented an overview of the evaluated algorithms with their corresponding test environment and number of training points. However these training points are not always the evaluation points used. In some cases, the training points are used for the *off-line* phase of the twofold solutions (e.g. fingerprinting).

Similar findings can be found in most research papers: the number of evaluation points is selected arbitrarily and information about the selection procedure is typically missing.

7.2.2 Evaluation procedure in competitions

The main goal of localization competitions is to objectively compare the performance of ILS in realistic conditions. It is known that the location of an evaluation point strongly influences the expected error distance. For example, evaluation points in corners or near walls often have larger error distances. However, not all solutions are impacted to the same level, thereby biasing the competition outcome towards those solutions that perform well in the selected evaluation conditions. These effect is strengthened in case a small number of evaluation points is used. As such, it is interesting to analyse the evaluation point selection procedures used in these competitions.

EvAAL (Evaluation Ambient Assisted Living (AAL) through Competitive Benchmarking) was the first organization that started evaluating and comparing multiple solutions in realistic conditions. Since their main focus is tracking, EvAAL benchmarks solutions use a path, but it is not clear which points on this path are evaluated. The EvAAL competition is held yearly at the Indoor Positioning & Indoor Navigation (IPIN) conference.

Since 2014, Microsoft started organizing indoor localization competitions yearly at the Information Processing in Sensor Networks (IPSN) conference. In contrast to EvAAL, Microsoft uses predefined static evaluation points to determine the performance accuracy of the evaluated solutions. In 2014 and 2015, 20 evaluation points were used. However, the latest competition in 2016 used only 15. Evaluation points are selected in advance, without following specific guidelines.

Finally, Evaluation of RF-based Indoor localization Solutions (EVARILOS) organized a competition with a special focus on the benchmarking methodology. The main focuses were the evaluation procedure in combination with the performance metrics and its definitions. During the competition, two different environments were used with each a minimum of 20 used evaluation points.

Based on this overview, it is clear that most competitions use 20 evaluation points or less to decide which solutions perform well. Since several competing solutions the past years obtained very similar accuracies, the question if this number is sufficiently high might impact future competition outcomes.

7.2.3 Anchor placement guidelines

Finally, the influence of evaluation points should not be confused with the influence of anchor point selection. [4, 11, 12] provide guidelines to place the anchor nodes of an indoor localization solutions whereby the best accuracy can be achieved. Their work is different from ours in the sense that this paper assumes that a system has already been deployed (using certain criteria which might include anchor point placement), but needs to be evaluated objectively after deployment using as few evaluation points as possible.

The summary, the above related work section describes papers which compare indoor localization solutions. Most paper lack information about the selected evaluation locations. An overview of the discussed solutions can be found in Table 7.1. To the best of our knowledge, not a single paper could be found that suggests guidelines that need to be applied in order to represent meaningful results.

7.3 Test set-up

The results discussed in this paper have been obtained using two test environments, two localization algorithms and multiple wireless technologies (Zigbee and Wi-Fi). Each time, one combination of a testbed, algorithm and technology was analysed to identify the impact of the number of evaluation points. Afterwards, the obtained results are verified using the other combinations. This section summarizes the used environments and algorithms in combination with the process of data gathering.

7.3.1 Test environments

Two different test environments have been used for the evaluations: w-iLab.t II (an industrial like environment) and Sint-Jozefskliniek (a hospital environment).

7.3.1.1 w-iLab.t II

The w-iLab.t II test facility is located in Zwijnaarde (Ghent, Belgium) above a cleanroom. The environment can be regarded as a challenging industrial environment that consists of a 66 m x 21 m area containing metal pipes and obstacles as shown in Figure 7.1. Since it is surrounded by metal walls, it

Table 7.1: Overview of papers and competitions with the used test environment and evaluation points.

Summary	Test Environment	Area [m^2]	# Evaluation Points	Ref
RESEARCH PAPERS				
Indoor Localization Sensing solutions	Typical buildings	[12, 8 625]	n.a.	[1]
Comparative evaluation #1	Building	44.8	21	[2]
Comparative evaluation #2	40th floor under construction	32.13	18	[2]
ORBIT	ORBIT Indoor Testbed	334.45	400	[3]
Optimal landmark placement	3rd floor of CoRE Building	1 486	286	[4]
High performance	Outdoor, open field	100 , 144	36, 49	[5]
COMPETITIONS				
EvAAL 2011 - 2016	Home environment	n.a.	Tracking a path	[6]
Microsoft Indoor Localization 2014	Large office	300	20	[7]
Microsoft Indoor Localization 2015	Exhibition hall	2000	20	[8]
Microsoft Indoor Localization 2016	Large office	465	15	[9]
EVARILOS 2014	Office	450	20	[10]

is a shielded environment without outside interference. Only weak external signals can be observed sporadically at the borders of the test area.

Figure 7.1: The w-iLab.t II testbed environment located in Zwijnaarde, near Ghent. The testbed is an open space containing metal pipes and obstacles causing multipath fading. Therefore, it is seen as a challenging environment.

In order to determine the influence of the selection of evaluation points, a fine mazed evaluation grid of of 2 m x 2 m is used, resulting in 203 evaluation points. A few points are missing in the grid, mainly in the middle of the test environment, due to metal obstacles and pipes that make it infeasible for the robot to drive to these locations. However, the total number of evaluation points is still a factor 10 larger than the number of evaluation points used in most scientific papers.

An advanced robot (based on a Roomba vacuum cleaner) drives through the entire testbed and collects measurement data at the predefined 203 evaluation points (Figure 7.2). At each evaluation point, the bot collected RSSI and Time of Arrival (ToA) values for 90 seconds using an STM-32W as a mobile device. The mobile device sends a unicast message to each available anchor node (in a loop). The anchors receiving such a message will reply with an Acknowledgement (ACK). The ACK Time of Arrival and RSSI data is stored in a Google Protocol buffer, making it is accessible to several applications and programming languages.

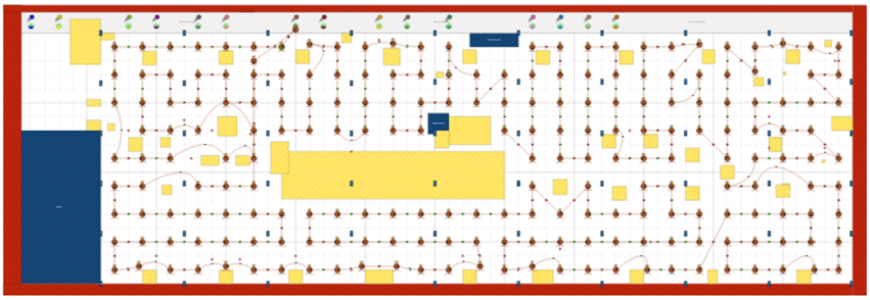


Figure 7.2: 203 evaluation points that were used in the w-iLab.t II test environment in Zwijnaarde. Yellow areas are not available due to pipes and other constructions. The blue areas indicate the technical equipment (e.g. dataracks) of the testbed. At these marked locations, it is not feasible to drive a robot.

7.3.1.2 Hospital environment

Another measurement campaign was executed in an actively used hospital environment (the Sint-Jozefskliniek hospital in Izegem, Belgium). The measurements were performed in the “surgical day hospital” ward, located in a new building on the first floor. The end of the section of the corridor was available for the experiments, while the rest of the ward was in “normal operation”, meaning patients and nurses were present and were walking around. The floor plan of the ward is depicted in Figure 7.3. Patient rooms 9, 10 and 11 were used for the evaluation. A dense evaluation grid of 1 m by 1 m was marked on the floor resulting in 73 evaluation locations where the position estimates were requested, i.e. more than double the number of evaluation points used in most scientific papers. Note that the grid was positioned in such a way that grid lines are 10 cm away from the wall. During the data collection, all doors were open.

7.3.2 Localization algorithms

The different evaluated localization algorithms use an off-line “raw” dataset as input, acquired using the EVARILOS benchmarking platform [13]. As such, multiple indoor localization algorithms can be applied to the same dataset.

7.3.2.1 Multilateration algorithm

Initially, a basic indoor localization approach is used. The algorithm processes the RSSI-values only. First, it averages all the values received per reachable anchor point. Next, the algorithm applies multilateration on all possible combinations with three different anchor nodes. Finally, it aver-

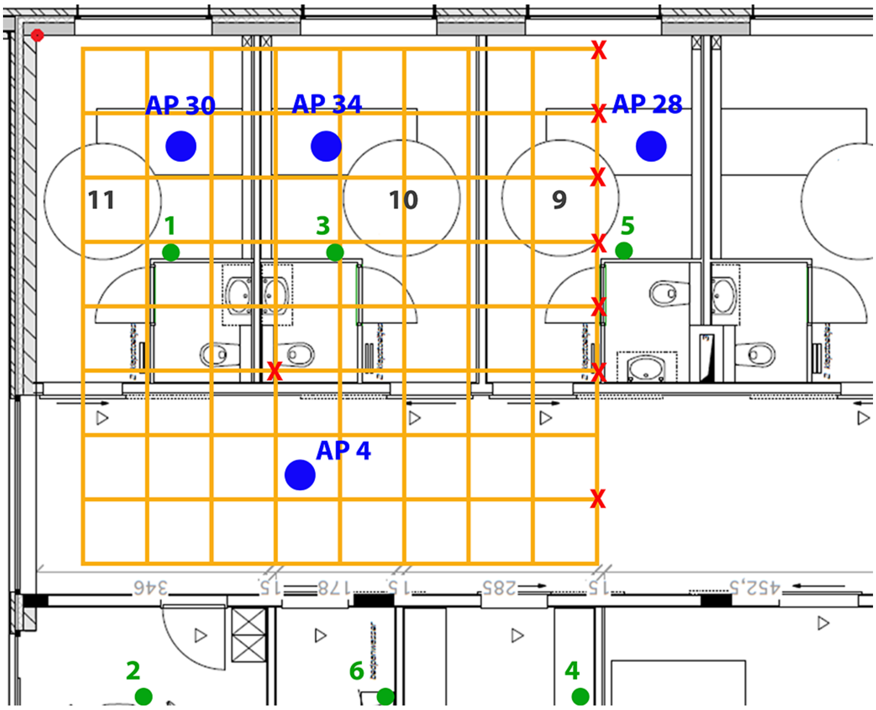


Figure 7.3: The floor plan of the hospital environment. The yellow grid represents the available evaluation points. The grid points with a red cross do not have measurement data, this results in 73 available evaluation points. Blue dots refer to the location of the Wi-Fi AP, the green ones represent the location of the Zigbee and BLE nodes.

ages all the x and y coordinates separately in order to return the final x and y coordinate of the estimated location.

7.3.2.2 Algorithm based on Viterbi

As an alternative, an advanced solution [14] is applied on the available datasets. The main principle is based on the Viterbi algorithm. Further, it uses semantic data to process and estimate the path that needs to be tracked. As expected, the accuracy results of this solution perform better than the solution described above. These results are further elaborated in Section 7.4.

7.4 Analysis of the behaviour of the error distances

This section describes the behaviour of the complete set of error distances. The accuracy results of both algorithms described in Section 7.3.2 are visualized in Figure 7.4, based on the w-iLab.t II dataset. As expected, the Viterbi solution achieved lower error distances (represented by lower error bars in the figure). In addition, Figure 7.4a demonstrates that, for both of the solutions, the point accuracy is typically worse at the edges in comparison with the center of the environment.

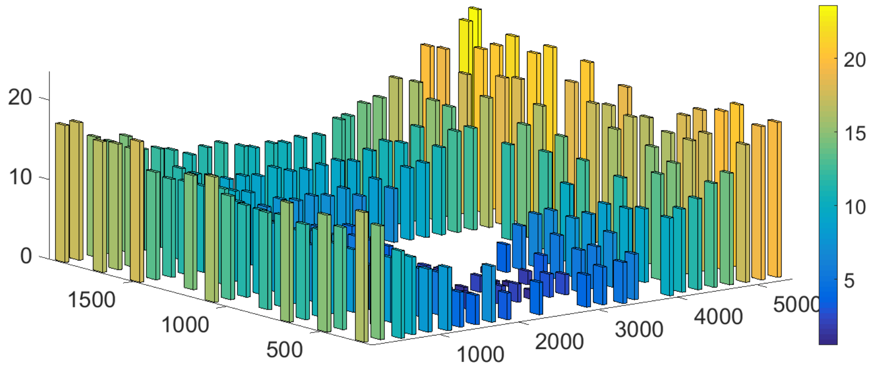
To gain better insight in the spread and distribution of the errors, multiple distributions are tested with these datasets in order to find the best matching one. The error distances of the first algorithm can be represented using a normal distribution ($X \sim N(\mu, \sigma)$ with $\mu = 10.46$m and $\sigma = 5.36$m). This is verified using a χ^2-test (chi-square test), that accepts the null hypothesis that the error distances are distributed normally with a 5% significance level. The histogram of the error distances is shown in Figure 7.5a in combination with the corresponding normal distribution.

For the the distribution of the error distances of the Viterbi algorithm, the χ^2-test rejects the null hypotheses. As such, these error distances are not normally distributed. Since the Viterbi algorithm has on average lower errors, the histogram bins are shifted to the left, with a cut-off error value at 0m, resulting in an asymmetrical distribution. The best matching distribution is the `Weibull` distribution where the probability density function is described as:

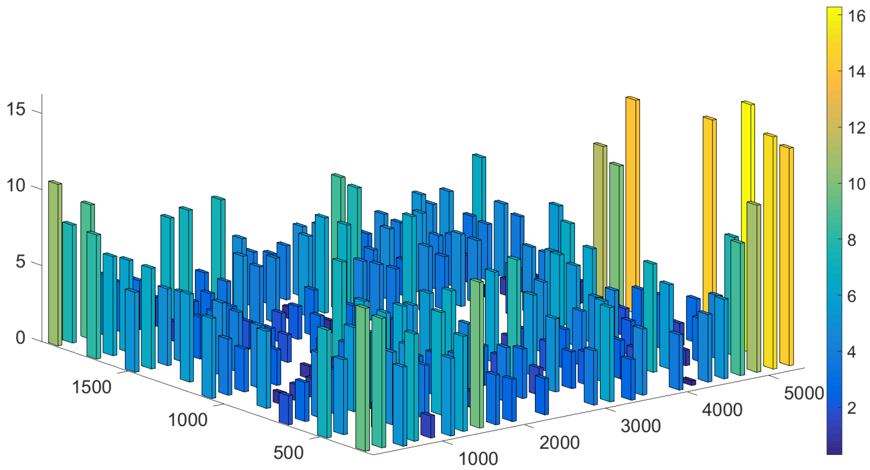
$$f(x;\lambda;k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0 \end{cases} \tag{7.1}$$

whereby $k > 0$ is the *shape* parameter and $\lambda > 0$ is the *scale* parameter. The fit is shown graphically in Figure 7.5b. The same Weibull distribution can also be used to fit the data from the first algorithm, making it a more generally applicable distribution.

In the remainder of this paper, the mean error value from this set of 203 evaluation points will be considered as the correct mean error of the evaluated solution. In other words, we assume that the mean value will not change further if additional evaluation points are included. This is a safe assumption since the used number of evaluation points is generally ten times higher than usual (cf. 203 instead of 20).

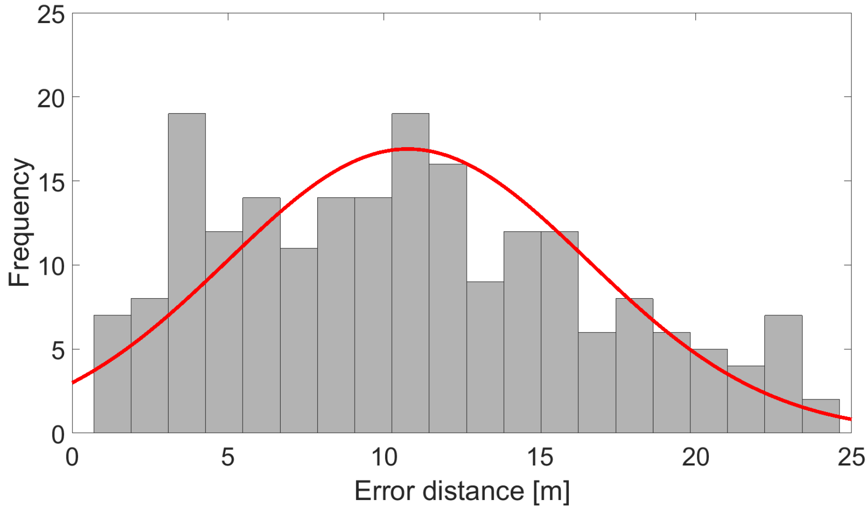


(a) Multilateration algorithm

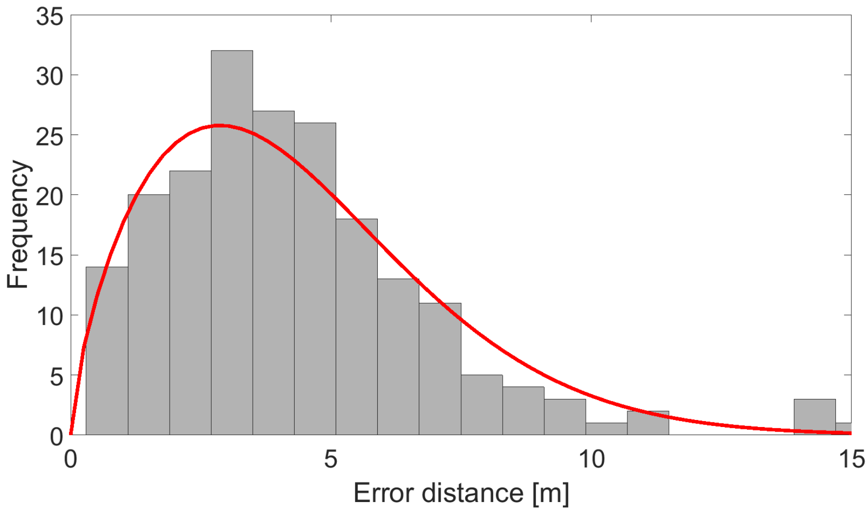


(b) Solution based on Viterbi

Figure 7.4: Visual representation of the error distances in a 3D bar plot. In 7.4a, it becomes clear error distances are higher at the edges of the test environment. Further, Figure 7.4b represents the best performing solution. The X and Y axes represent the coordinates of the testbed (in cm) whilst the Z axis shows the error distance (in m).



(a) Multilateration algorithm: normal distribution



(b) Solution based on Viterbi: Weibull distribution

Figure 7.5: Histogram of both the error distances from 203 evaluation points, together with the matching normal or Weibull distribution.

Takeaways:

- The distribution of the error distances at different evaluation points can be fitted using a Weibull distribution.
- The distribution of the error distances from localization algorithms with large error distances can be represented using a normal distribution ($X \sim N(\mu, \sigma)$), which is an easier distribution for use in statistics.

7.5 Impact of the number of used evaluation points

This section discusses the impact of the number of used evaluation points on the mean and standard deviation of the estimated accuracy. As discussed before: an evaluation point is a location in the testbed whereof the exact coordinates are known. The System Under Test (SUT) must calculate and estimate the coordinates of this location. The error distance for an evaluation point is defined by the euclidean distance between the estimated coordinates and the true coordinates.

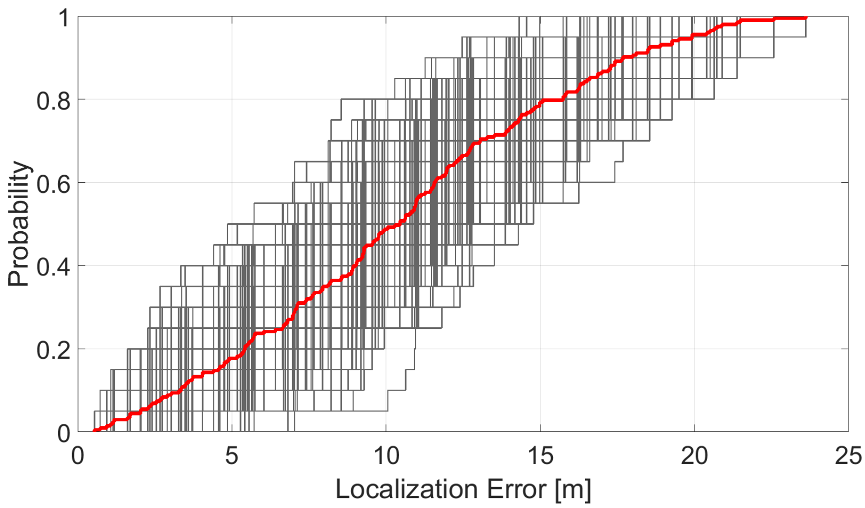


Figure 7.6: Multiple CDF plots for a subset of 20 evaluation points. Grey lines demonstrate the possible variance of the plot whilst the red line is the final CDF when all available evaluation points are used.

To demonstrate the impact of using different evaluation points, Figure 7.6 plots different CDF plots that can be obtained when using only 20 evaluation points out of the 203 evaluation points. CDF plots are often used to evaluate the effectiveness of a localization solution. However, as Figure 7.6 demonstrates, the behaviour of the mean error distance can vary significantly depending on which evaluation points are used.

This section provides a thorough analysis of the behaviour of the estimated accuracy for different number of evaluation points. Figure 7.7 shows a flowchart of the steps executed during the analysis of the results. The process is divided in two steps. Step IV processes the statistics of the entire dataset, these results are already presented in Section 7.4. Step V calculates the statistics of smaller subsets of n evaluation points. For each n, 200 randomly selected subsets ($n = [2, 202]$) are used to calculate statistics. This results in a matrix of data for each statistical parameter (Figure 7.11).

7.5.1 Multilateration algorithm in w-iLab.t II

To determine the impact of the used number of evaluation points, it is important to analyse the spread of the estimated mean error distance for different numbers n of evaluation points. χ^2-tests confirm the spread is normally distributed. Appendix 7.A describes in detail how the data is processed to obtain the results below.

Figure 7.8 shows the spread of the estimated error as a function of the number of used evaluation points. As the number of evaluation points increases, the variance of estimated mean error distance decreases. However, it is worth noting that even for large numbers of evaluation points, the variance remains quite high. As a result, even when using e.g. 60 evaluation points, in this case the estimated mean error distance can vary between 9.02 and 12.06 meters. Luckily, as the number of evaluation points increases, the probability that a set of evaluation points is selected that results in such extreme values also decreases.

Takeaway: For a fixed number of n evaluation points, multiple random subsets can be defined. The resulting mean error distances are normally distributed with decreasing variance for increasing values of n.

A high standard deviation means the likelihood of an atypical subset becomes higher. Intuitively, it makes sense that, if the number of used evaluation points is low, the variation of the standard deviation will be high. Next, we analyse in more detail the behaviour of the standard deviation (also represented as σ), which can be used as a representative of the confidence a researcher can assign to the reported mean accuracy. Assuming the error distance is defined as e and the collection of evaluation points is defined by the variable X in combination with the error distances, i.e. $e = f(X)$. When N equals the total number of used evaluation points

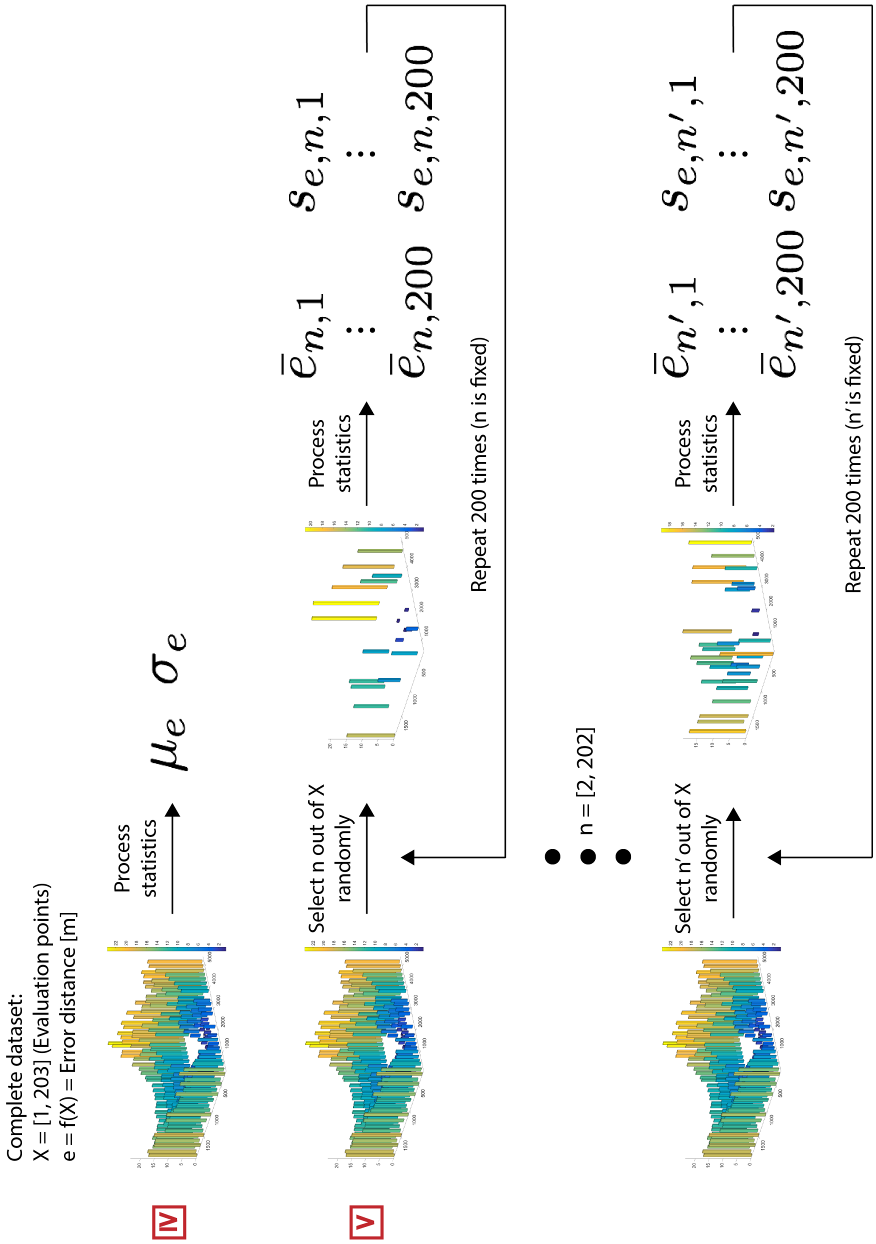


Figure 7.7: The flowchart of the data-processing. Section 7.4 discusses the statistics retrieved by following step IV*, (i.e. the behaviour of all the 203 error distances). Section 7.5 present the results of the smaller subsets of n evaluation points. These subsets are created following step* V.

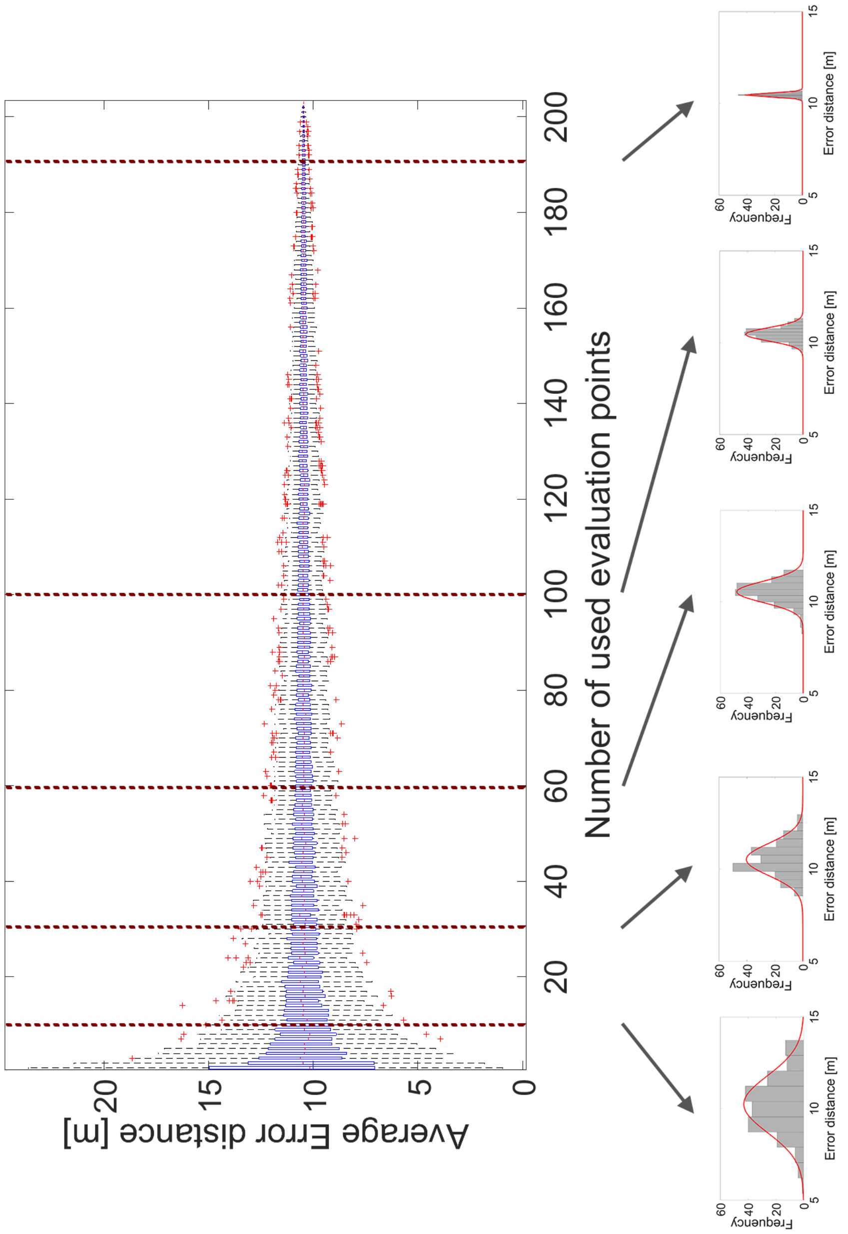


Figure 7.8: Boxplots representing the spread of the mean error distances with different subsets of evaluation points. Below the boxplots, histograms visualize the spread of respectively n = 10, 30, 60, 100 and 190 evaluation points.

(i.e. 203), the standard deviation of the error distances from the complete dataset can be calculated using the Equation 7.2.

$$\sigma_e = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|e_i - \mu_e|^2} \text{ with } \mu_e = \frac{1}{N}\sum_{i=1}^{N} e_i \tag{7.2}$$

The above formula can be used when the "complete set" of evaluation points is available, in this case with $N = 203$. However, when smaller evaluation point sets are used, the standard deviation σ_e is estimated by examining a random sample taken from the complete dataset and computing a statistic of the sample. This statistic is called an estimator, namely the `sample standard deviation` s_e. In order to calculate the value of this estimator, Bessel's correction [15] needs to be applied. It is the use of $N-1$ instead of N in Equation 7.2.

$$\sigma_e \approx s_e = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}|e_i - \bar{e}|^2} \text{ with } \bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i \tag{7.3}$$

Whereby N is the number of samples and $\bar{e}$ is the sample mean. The behaviour of the sample standard deviation s_e becomes more stable as the number of evaluation points increases. It is comparable with the behaviour of the sample mean value $\bar{e}$.

For determining a suitable number of evaluation points, it is important to consider the standard deviation of the sample mean values $\sigma_{\bar{e}}$ (see Appendix 7.A step $\boxed{\text{B}}$). The relation between the standard deviation of the mean value $\bar{e}$ and the number of used evaluation points (n) is depicted in Figure 7.9. The decreasing curve confirms the importance of a decent number of evaluation points. Each additional evaluation point results in a lower standard deviation $\sigma_{\bar{e}}$. In other words: the sample mean values are more stable and have less variation if the number of evaluation points increases. However, at certain level, the "cost" to add an additional evaluation point will result in each time smaller marginal improvements. This is also known as `the law of diminishing returns`.

Based on the central limit theorem, the standard deviation of the mean values can be calculated as follows:

$$\sigma_{\bar{e}} = \frac{\sigma_e}{\sqrt{N}} \tag{7.4}$$

whereby σ_e is the true standard deviation of the entire dataset. In realistic evaluation set-ups, it is not feasible to calculate these statistics since a complete dataset is lacking. In this case, an estimator can be used for $\sigma_{\bar{e}}$.

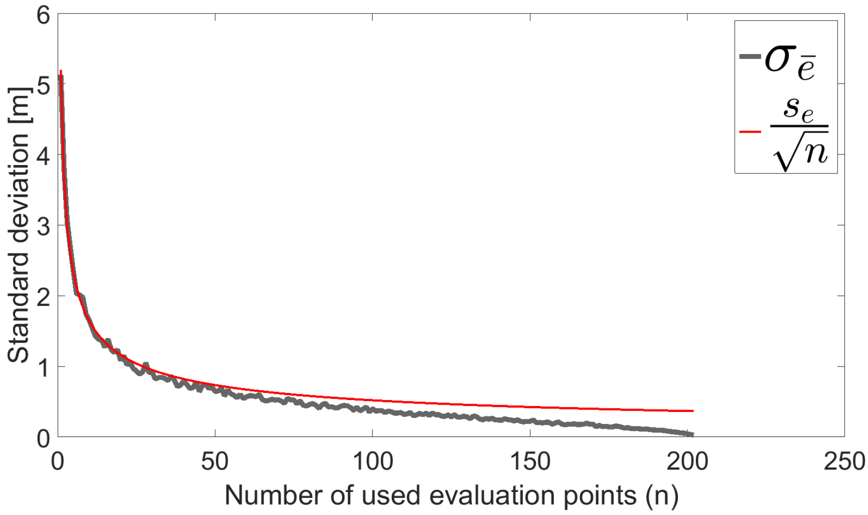


Figure 7.9: Standard deviation of the mean values of the error distances $\sigma_{\bar{e}}$ are represented by the grey curve. The estimator SEM is visualized by the red curve and is an efficient estimation of the standard deviation values when the number of evaluation points (n) is lower than the total number of evaluation points N (e.g. 203).

This estimator is known as the `standard error of the mean (SEM)`:

$$\sigma_{\bar{e}} \approx SEM = \frac{s_e}{\sqrt{N}} \tag{7.5}$$

In conclusion, it is possible to estimate the behaviour of the standard deviation of the mean values for each value of n ($n = [0, N]$) by only using the sample standard deviation of a single sample. This estimation is shown in Figure 7.9 as well (the red curve), with s_e being the standard deviation of a subset of 30 evaluation points. It is also possible to use lower number of samples to calculate s_e, thereby reducing the number of required evaluation points to a more feasible number. In this case, the estimator will have higher shifts [16].

For large number of evaluation points, i.e. more than 100 in Figure 7.9, the estimator does not follow the standard deviation any more, since it assumes that the sample size is much smaller than the size of the complete dataset. As such, estimator (7.5) is only appropriate when the complete set is considered as an infinite set of evaluation points or if the complete set is unknown. To obtain a better estimation also for these subsets of evaluation points, the estimator must be corrected by multiplying a `finite population correction (FPC)` [17].

$$\sigma_{\bar{e}} \approx SEM = \frac{s_e}{\sqrt{N}} * \sqrt{\frac{N-n}{N-1}} \tag{7.6}$$

Figure 7.10 demonstrates the effect of the finite population correction. The estimator SEM almost exactly fits the calculated standard deviations using only two parameters: the standard deviation form a single sample and the total number of evaluation points. In the cases whereby the total number of available sample point is unknown, using only one single parameter will be sufficient.

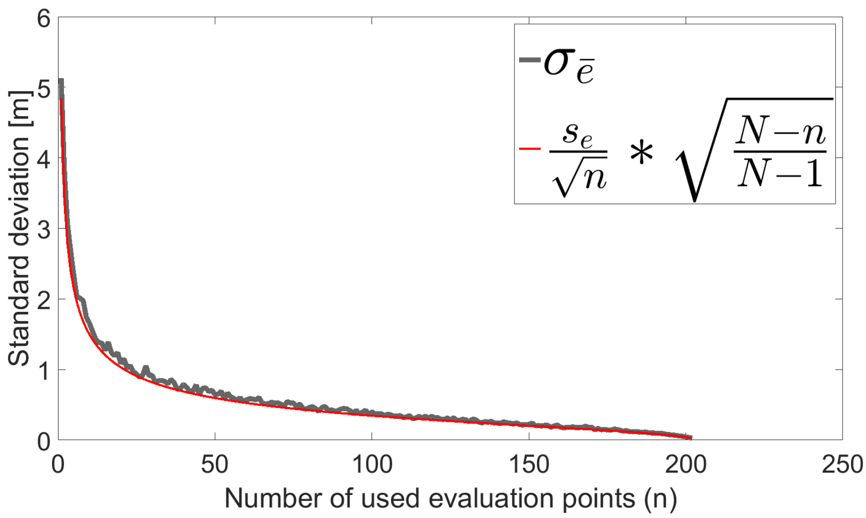


Figure 7.10: Standard deviation of the mean values of the error distances $\sigma_{\bar{e}}$ in combination with the estimator SEM whereby the finite population correction is applied.

Takeaway:

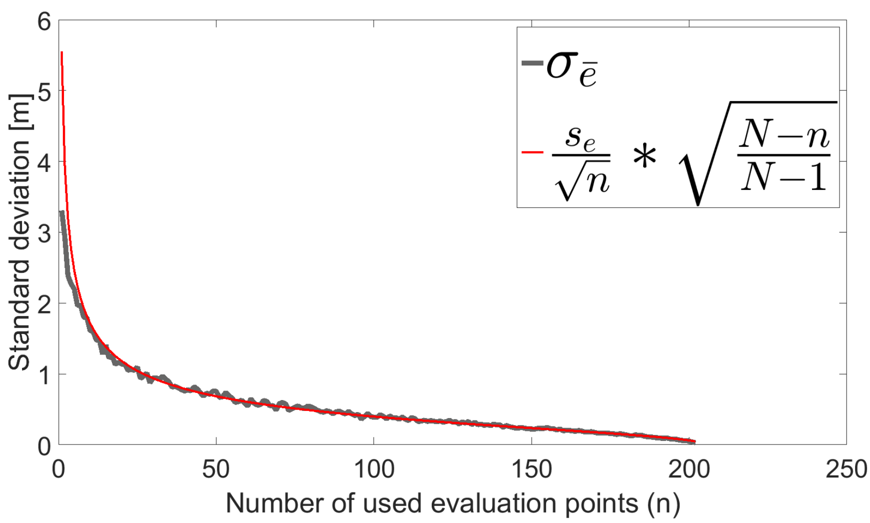
- The standard deviation of the mean error distances is an indicator of the confidence one can have about the estimated mean error.
- The standard deviation of the mean values can be estimated for different values of n by using the standard deviation of the error distances s_e of one single subset of evaluation points.
- The approximation of the standard error function is sufficient to give the experimenter an idea about the true standard deviation, even when using low values for n. As such, an experiment with a chosen value for n can be used to identify how many additional evaluation points are required to obtain the mean error distance within certain variance levels.

7.5.2 Verification

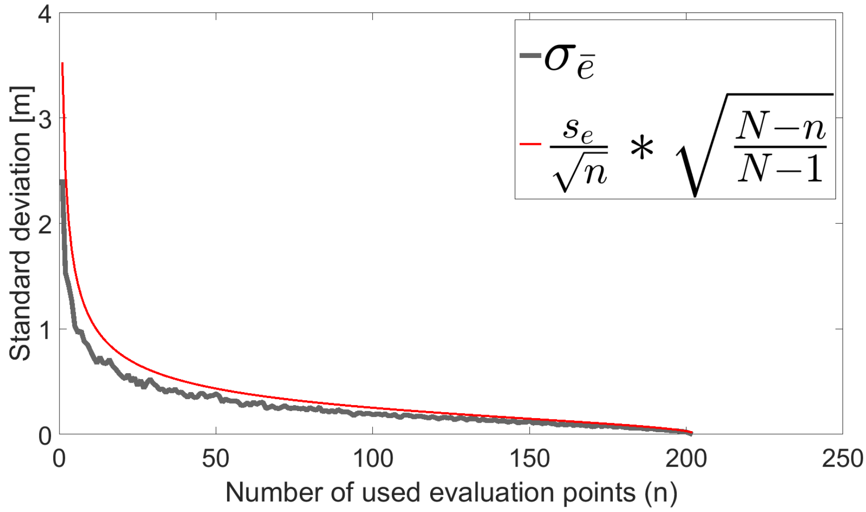
This section verifies if the conclusions above are also applicable for other conditions, by verifying our findings for (i) a different localization algorithm (Viterbi), (ii) a different environment (hospital) and (iii) a different technology (IEEE 802.11 Wi-Fi). The relation between the estimator $\sigma_{\bar{e}}$ and $\frac{s_e}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$ (based on a subset of 30 evaluation points) and the real behaviour of the standard deviation is shown in Figure 7.10 for each configuration. All the standard deviations and mean values (of the complete dataset) are listed in Table 7.2.

Table 7.2: Summary of the statistical results. All available mean μ and standard deviations σ are summarized in this table. These numbers are calculated using the entire dataset.

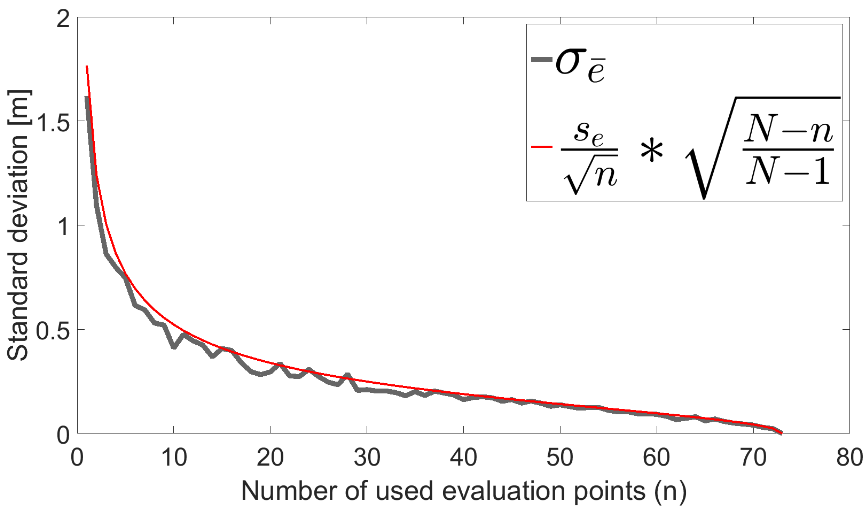
Algorithm	Technology	w-iLab.t II [m]		Hospital [m]	
		μ_e	σ_e	μ_e	σ_e
Multi-lateration	Wi-Fi	10.77	5.79	2.68	1.37
	Zigbee	10.46	5.36	3.08	1.64
	BLE	n.a.	n.a.	3.12	1.51
Viterbi	Zigbee	4.35	2.76	n.a.	n.a.



(a) Technology: Wi-Fi



(b) Algorithm: Viterbi



(c) Environment: Hospital

Figure 7.10: Plot of the calculated standard deviation (based on 30 evaluation points) of the mean values with the standard error function. In all cases, s_e is the standard deviation of 30 error distances, selected randomly in the environment.

Takeaway:

- The formulas for calculating the expected standard deviation for different number of evaluation points are valid also for other localization algorithms, environments and wireless technologies.

7.6 General guidelines

In this section, practical guidelines to answer the question "How many evaluation points are essential in order to achieve an acceptable mean value which is close enough to the actual mean?" are formulated. These guidelines do not have the intention to impose a minimum number of evaluation points, only to allow a researcher to identify how many experiments are required to identify the mean error distances within predefined variance bounds.

The *five-step* guide:

Step 1 Select randomly as independent as possible $N = 20$ evaluation points in the chosen test environment.

Step 2 Calculate the mean value $\bar{e}_N$ and the standard deviation $s_{e,N}$ of the dataset.

Step 3 Calculate the current confidence interval based on these statistics:

$$\left[\bar{e}_N - 1.96 * \frac{s_{e,N}}{\sqrt{N}}, \bar{e}_N + 1.96 * \frac{s_{e,N}}{\sqrt{N}}\right] \tag{7.7}$$

Step 4 Define the desired confidence bounds CB ($\bar{e}_N \pm CB$) and calculate the amount of evaluation points N' needed:

$$N' = \left(\frac{1.96 * s_{e,N}}{CB}\right)^2 \tag{7.8}$$

Step 5 Go back to **Step 1** with $N = N'$ and verify the confidence bounds are changed.

To demonstrate the usefulness of these guidelines, they have been applied to different combinations of localization algorithms, wireless technologies and environments in Table 7.3. In this verification, a confidence bound of $1\ m$ was chosen. The number of required evaluation points vary from 20 up to 109 evaluation points. The outcomes also demonstrate that is meaningless to define general minimum requirement regarding the number of required evaluation points. Instead, the guidelines are used to define the acceptable confidence interval, which will determine the number of necessary evaluation points. Not all *error distances* are found within the confidence interval. However, the **mean value** of the error distances is within the confidence bounds, which is typically the value that is reported in scientific literature.

Takeaway:

- The stability of the standard deviation depends on multiple factors like algorithm, technology and test environment. Therefore, the number of necessary evaluation points is also dependent on these factors.
- Guidelines are provided to determine the number of necessary evaluation points in order to obtain the mean error within a desired confidence interval.

7.7 Conclusions

This paper addresses the lack of a thorough performance evaluation in published research work relating to novel indoor localization solutions. The number and location of evaluation points is selected without any reasonable study. Each solution is evaluated in its own environment, using self-selected evaluation points which can bias these results.

Since evaluating an indoor localization solution is often time consuming, many researchers choose to limit the number of evaluation points. However, as demonstrated in this paper, the location and the number of evaluation points can influence the published performance results of the SUT. In this paper, a study on the behaviour of the error distances is provided. It is shown that a Weibull distribution appears to be the best fitting distribution for multiple datasets of error distances. If the algorithm has larger error distances, this distribution can be simplified by using a normal distribution.

Since most researchers report the mean error distances of evaluated solutions, it is useful to have an indication of the correctness of this estimate. Towards this end, the paper has shown that it is possible to estimate the standard deviation of the mean value for a fixed number of evaluation points. Based on this standard error, a confidence interval for the mean error distance can be calculated. Simple guidelines are provided that help researchers find the optimal number of evaluation points in their current

Table 7.3: Applying the guidelines on the multiple datasets of error distances. The number of necessary evaluation points depends on the combination of the selected algorithm, technology and environment. It also depends on the desired confidence bounds of the mean value. In this example, a confidence bound of one meter was chosen.

Algorithm	Multilateration	Multilateration	Viterbi algorithm	Multilateration
Technology	Zigbee technology	Wi-Fi technology	Zigbee technology	Zigbee technology
Environment	w-iLab.t II	w-iLab.t II	w-iLab.t II	Hospital
Step 1	$N = 20$	$N = 20$	$N = 20$	$N = 20$
Step 2	$\bar{e}_{20} = 9.88\ m$	$\bar{e}_{20} = 10.81\ m$	$\bar{e}_{20} = 3.89\ m$	$\bar{e}_{20} = 2.89\ m$
	$s_{e,20} = 4.63\ m$	$s_{e,20} = 5.32\ m$	$s_{e,20} = 2.28\ m$	$s_{e,20} = 1.79\ m$
Step 3	$[9.88 \pm 2.03]\ m$	$[10.81 \pm 2.33]\ m$	$[3.89 \pm 1.00]\ m$	$[2.89 \pm 0.78]\ m$
	$[7.85, 11.91]\ m$	$[8.48, 13.14]\ m$	$[2.89, 4.89]\ m$	$[2.11, 3.67]\ m$
Step 4	Desired $CB = 1\ m$	Desired $CB = 1\ m$	Desired $CB = 1\ m$	Desired $CB = 1\ m$
	$\Rightarrow \boldsymbol{N' = 82}$	$\Rightarrow \boldsymbol{N' = 109}$	$\Rightarrow \boldsymbol{N' = 20}$	$\Rightarrow \boldsymbol{N' = 20}$
Step 5	$\bar{e}_{82} = 11.17\ m$	$\bar{e}_{109} = 10.45\ m$	$\bar{e}_{20} = 3.89\ m$	$\bar{e}_{20} = 2.89\ m$
	$s_{e,82} = 5.40\ m$	$s_{e,109} = 5.63\ m$	$s_{e,20} = 2.28\ m$	$s_{e,20} = 1.79\ m$
	$[11.17 \pm \mathbf{1.16}]\ m$	$[10.45 \pm \mathbf{1.05}]\ m$	$[3.89 \pm \mathbf{1.00}]\ m$	$[2.89 \pm \mathbf{0.78}]\ m$
	$\mathbf{[10.01, 12.33]}\ \boldsymbol{m}$	$\mathbf{[9.40, 11.50]}\ \boldsymbol{m}$	$\mathbf{[2.89, 4.89]}\ \boldsymbol{m}$	$\mathbf{[2.11, 3.67]}\ \boldsymbol{m}$
Mean value $\bar{e}$	$10.46\ m$	$10.77\ m$	$4.35\ m$	$3.08\ m$

situation. As such, the outcomes allow finding the optimal number of evaluation points in order to obtain the desired confidence interval for the mean value of the error distances. Finally, all results are experimentally validated using multiple localization algorithms, multiple wireless technologies and in different environments, demonstrating the generic nature of the obtained results.

7.A Calculation of the mean and variance

This section describes how the data traces are processed to obtain the mean and variance of the estimated location error for different subsets of n evaluation points. Figure 7.8 is achieved by processing the data using Trace $\boxed{\text{V}}$ as depicted in Figure 7.7. This results in a matrix that consists of mean values $\bar{e}_{i,j}$. The i-th column represents the number of repetition whilst the j-th column stands for the number of used evaluation points. Representing the vertical mean values ([$\bar{e}_{1,n}$,$\bar{e}_{200,n}$]) into a boxplot for each value of n is visualized in Figure 7.8. This process is visualized in Figure 7.11 (Part $\boxed{\text{A}}$).

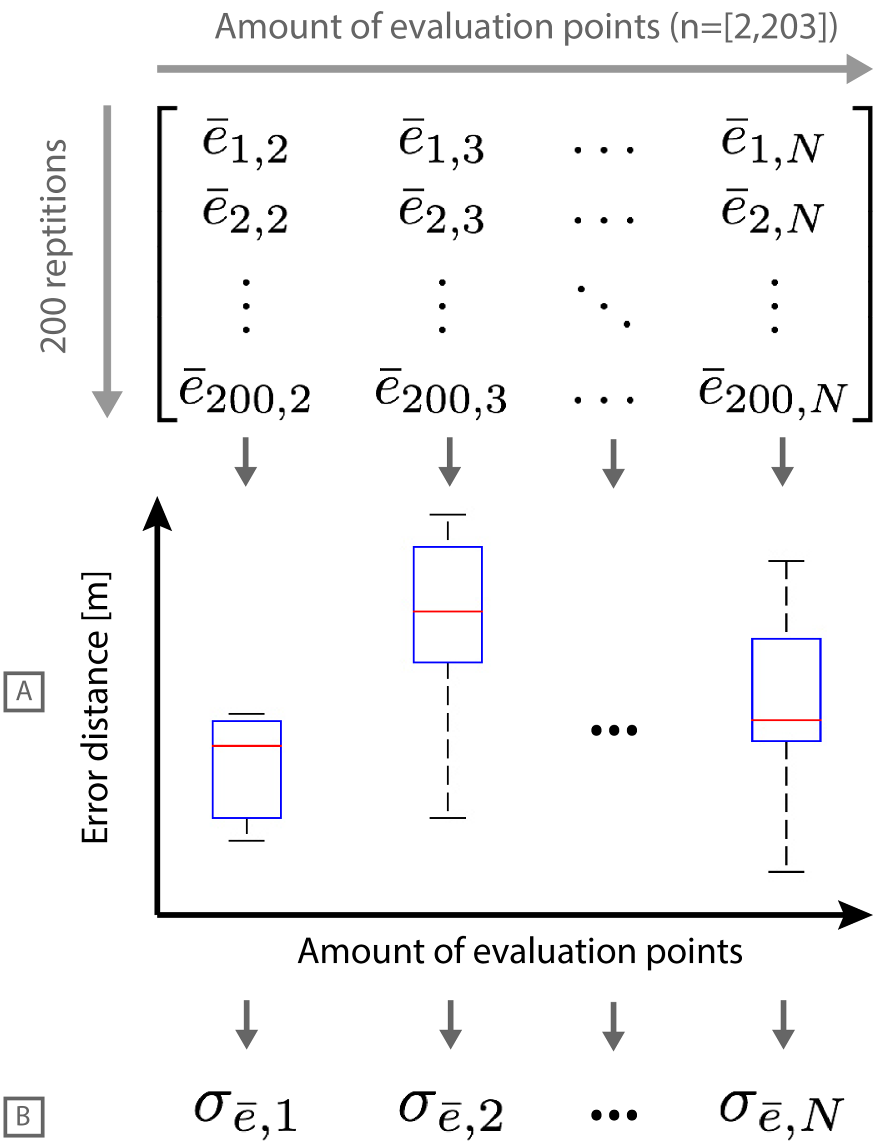


Figure 7.11: Processing the statistics according to the flow in Figure 7.7, a matrix of mean values $\bar{e}$ is achieved. Using this matrix, Figure 7.8 (A) and Figure 7.9 (B) have been established.

References

[1] N. Li and B. Becerik-Gerber. *Performance-based evaluation of RFID-based indoor location sensing solutions for the built environment*. Advanced Engineering Informatics, 25(3):535–546, 2011.

[2] X. Luo, W. J. OBrien, and C. L. Julien. *Comparative evaluation of Received Signal-Strength Index (RSSI) based indoor localization techniques for construction jobsites*. Advanced Engineering Informatics, 25(2):355–363, 2011.

[3] G. Chandrasekaran, M. A. Ergin, J. Yang, S. Liu, Y. Chen, M. Gruteser, and R. P. Martin. *Empirical evaluation of the limits on localization using signal strength*. In Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON'09. 6th Annual IEEE Communications Society Conference on, pages 1–9. IEEE, 2009.

[4] Y. Chen, J.-A. Francisco, W. Trappe, and R. P. Martin. *A practical approach to landmark deployment for indoor localization*. In Sensor and Ad Hoc Communications and Networks, 2006. SECON'06. 2006 3rd Annual IEEE Communications Society on, volume 1, pages 365–373. IEEE, 2006.

[5] M. I Marks and E. Niewiadomska-Szynkiewicz. *High Performance Wireless Sensor Network Localization System*. 2013.

[6] J. A. Álvarez-García, P. Barsocchi, S. Chessa, and D. Salvi. *Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 EvAAL competition*. Journal of Ambient Intelligence and Smart Environments, 5(1):119–132, 2013.

[7] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, S. Sen, and V. Handziski. *Microsoft indoor localization competition: Experiences and lessons learned*. GetMobile: Mobile Computing and Communications, 18(4):24–31, 2015.

[8] Microsoft. *Microsoft indoor localization competition 2015*. http://research.microsoft.com/en-us/events/indoorloccompetition2015/, 2015. [Online; accessed 18-April-2016].

[9] Microsoft . *Microsoft indoor localization competition 2016*. http://research.microsoft.com/en-us/events/msindoorloccompetition2016/, 2016. [Online; accessed 18-April-2016].

[10] F. Lemic, V. Handziski, A. Wolisz, T. Constambeys, C. Laoudias, S. Adler, S. Schmitt, and Y. Yang. *Experimental evaluation of RF-based indoor localization algorithms under RF interference*. In Localization and GNSS (ICL-GNSS), 2015 International Conference on, pages 1–8. IEEE, 2015.

[11] T. Kunz and B. Tatham. *Localization in Wireless Sensor Networks and Anchor Placement*. Journal of Sensor and Actuator Networks, 1(1):36, 2012.

[12] K. Balac, M. Prevostini, and M. Malek. *Optimizing Sensor Nodes Placement for Fault-tolerant Trilateration-based Localization*. In Dependable Computing (PRDC), 2015 IEEE 21st Pacific Rim International Symposium on, pages 109–118. IEEE, 2015.

[13] T. Van Haute, E. De Poorter, F. Lemic, V. Handziski, N. Wirstrom, T. Voigt, A. Wolisz, and I. Moerman. *Platform for benchmarking of RF-based indoor localization solutions*. Communications Magazine, IEEE, 53(9):126–133, 2015.

[14] J. Trogh, D. Plets, L. Martens, and W. Joseph. *Advanced Real-Time Indoor Tracking Based on the Viterbi Algorithm and Semantic Data*. International Journal of Distributed Sensor Networks, 2015, 2015.

[15] W. J. Reichmann. *Use and abuse of statistics*. 1964.

[16] S. S. Kar and A. Ramalingam. *IS 30 THE MAGIC NUMBER? ISSUES IN SAMPLE SIZE ESTIMATION*. National Journal of Community Medicine, 4(1), 2013.

[17] L. Isserlis. *On the value of a mean as calculated from a sample*. Journal of the Royal Statistical Society, 81(1):75–81, 1918.

8 Conclusion

"The greatest challenge to any thinker is stating the problem in a way that will allow a solution."

–Bertrand Russell (1872 - 1970)

In the age of automation, the ability to locate persons and assets in indoor environments has become increasingly important for a rising number of application domains. With the emergence of Global Positioning Systems (GPSs), the performance of outdoor positioning has become excellent, but many applications require seamless positioning capabilities in all environments. Therefore, indoor positioning has become a focus of research and development during the past decade. The diversity of different technological solutions for indoor positioning and navigation reflects that almost any signal/sensor technique can be exploited for this purpose.

The main goal of this PhD research was to orchestrate these solutions by defining a clear benchmarking methodology whereby a comparison between multiple solutions became possible. This methodology must act as a strong message to every researcher who publishes performance results concerning Indoor Positioning Systems (IPSs). The remainder of this final chapter summarizes the most important contributions, the main conclusions of the performed research and suggest different directions for future work.

8.1 Summary and conclusions

The main research question addressed in this dissertation was "How to evaluate and benchmark multiple Radio-Frequency (RF)-based localization solutions in heterogeneous environments in such a way a full comparison becomes possible?". Multiple aspects (like metrics, evaluation points, environments, etc.) come in the picture in order to formulate an adequate answer.

Although it was not the main focus of this PhD research, a first contribution was the **optimization of a Time of Arrival (ToA)-based solution** in a harsh industrial environment. Multipath fading has a great influence on the ranging measurements and thus on the estimated location of the mobile node. Calibrating and filtering outliers can definitely improve the accuracy results of solutions applied in environments where a lot of multipath fading may occur. The outcome of this work revealed room for improvement concerning the comparison of different solutions mutually. Up till now, researchers only publish point accuracy results achieved in their own environment making a fair comparison impossible.

To remedy this, a second contribution was the **design of a benchmarking methodology**. This methodology is summarized in a handbook which can be used as a guide for researchers evaluating their own IPS. A first task in this contribution was defining the metrics. From the perspective of the end-user: which parameters will influence my decision in selecting the "optimal" solution? The answer will be different depending on the needs of the end-user. As such, multiple parameters (like latency, energy consumption, scalability, installation time, interference robustness, etc.) need to be taken into account. Defining metrics is one thing, measuring them correctly is another matter. Therefore, benchmarking scenarios try to describe the entire process to evaluate a solution as objective as possible. Matters like the type of environment, the selection of evaluation points, how to generate artificial interference, etc. are described in detail. Finally, the handbook is aligned with the upcoming ISO/IEC 18305 standard: "Test and Evaluation of Localization and Tracking Systems". It supports all possible techniques of localization, in contrast to our handbook which is focussed on RF-based solutions.

Researchers who want to apply this benchmarking methodology encounter difficulties since the methodology states the chosen environment can strongly influence the performance results. As a solution, a third contribution was the **development of a web-based benchmarking platform** which allows researchers to evaluate their solution according to the proposed benchmarking methodology. Custom defined datasets can be acquired and reused for evaluation of multiple localization algorithms. Due to the public availability of these datasets, every researcher in the indoor localization community can download and use these datasets as an input for their localization algorithm. Since the input is identical for different solutions, an objective comparison is possible. Besides, the platform also supports visualization tools and statistical parameters in order to give a profound insight in the behaviour of the solution. Finally, this platform provides the opportunity to define an application profile whereby the related scores for each solution are calculated. In other words, we realized the possibility to translate an IPS into a final "score" based on multiple metrics with custom defined weights and boundaries. This was a major step in making indoor localizations comparable with each other.

A final contribution in this dissertation was a **profound investigation of the amount of evaluation points**. "How many evaluation points are necessary in order to represent a reliable mean value of the performance metrics?" was the main research question in this contribution. In current literature, every IPS is evaluated based on a certain experiment set-up. If the amount of evaluation points is mentioned in the set-up description, it fluctuates between 10 and 30 measurement points without substantiating why this amount is sufficient to represent a reliable result. The benchmarking methodology does a first attempt by defining minimum requirements (use at least 20 evaluation points, draw a grid and apply Latin hypercube sampling, etc.) however it does not formulate an answer on the research question stated above. Therefore, the final contribution defines a correlation between the amount of evaluation points and the confidence bounds of the mean error distance. In other words, the desired confidence interval will determine the amount of evaluation points that needs to be used.

The contributions in this PhD research work have proven to be an added value in the community of indoor localization. The awareness of the fact a more advanced evaluation methodology was lacking, has been strengthened. Researchers will pay more attention in processing the performance results than in the past. This will result in a better and more objective comparability between the different solutions.

8.2 Outlook

This work presented a benchmarking philosophy making comparability possible between different indoor localization solutions. This work is an initial head start to awake the awareness of the need of profound evaluation. However due to the continuous advancement of technology, new perspectives for future research arises.

A first direction for future work is **the integration of smartphone based localization** applications in the entire benchmarking methodology and platform. The capabilities of the smartphone are still not at their limits. The number of implemented features is still rising. New technologies like the 5^{th} Generation (5G) or Near Field Communication (NFC) can create new opportunities. Further, the Inertial Measurement Unit (IMU) (a combination of accelerometers, gyroscopes and sometimes magnetometers) can "measure" specific forces executed by a human body. This IMU cannot be influenced by interference, making it a valuable contribution in the estimation of the smartphone's location. Moreover, the integration of smartphone based solutions can go tremendously fast since nowadays everyone owns a smartphone. However, allowing smartphone based applications in the benchmarking methodology requires multiple adaptations: the raw datasets have to be recollect using a smartphone and extended with additional data (e.g. IMU-data). It will be challenging to collect realistic IMU-data in combination with repeatability since robots cannot simulate the behaviour of a person holding a smartphone in his hand or pocket. Further, the methodology has to be revised as well since the smartphone's position (to the ear, in the hand, in the pocket, etc.) also needs to be defined in the scenario descriptions.

The integration of smartphone based solutions inevitably involves another direction for future work: **the integration of indoor *tracking* solutions**. Currently, the benchmarking methodology is limited to RF-based indoor *positioning* solutions. Though, if smartphone based applications will be integrated, the benchmarking methodology must be able to cope with tracking as well. Evaluation points need to be translated into evaluation tracks taking into account distance, speed, cornering, direction, breaks, stairs, etc. The raw data collection will be more labour-intensive since it must be recorded manually. Metrics like point and room accuracy no longer suffice and need to be update in such a way tracks can be comparable.

Finally, a third direction for future work is **the determination of the evaluation points**. In Chapter 7, an initial guide for selecting evaluation points is given. However, this guide focusses on the amount of necessary evaluation points. Another important and still unanswered question is "Where to select the evaluation points?" Are minimum requirements necessary like: all the corners of the test environments need to be included or at least X evaluation points must be selected close to a wall. Does equally spreading the evaluation points improve the mean value or is picking them randomly sufficient? Answers to those questions will improve the evaluation of the indoor localization solutions. Those answers are currently lacking in scientific literature.

A

Hybrid Indoor Localization Solution Using a Generic Architectural Framework for Sparse Distributed Wireless Sensor Networks

This chapter presents a hybrid indoor localization solution based on a "generic framework". It is based on a combination between proximity and Received Signal Strength Indicator (RSSI) multilateration acquired using Zigbee technology. Thanks to the technology abstraction layer in the generic framework, other technologies can be added or changed easily without modifying the positioning calculator. This hybrid solution is utilized in the previous chapters as one of the solutions in the comparison studies.

⋆ ⋆ ⋆

T. Van Haute, J. Rossey, P. Becue, E. De Poorter, I. Moerman, P. Demeester.

Published in the proceedings of Federated Conference on Computer Science and Information Systems (FedCSIS), September 2014.

Abstract Indoor localization and navigation using wireless sensor networks is still a big challenge if expensive sensor nodes are not involved. Previous research has shown that in a sparse distributed sensor network the error distance is way too high. Even room accuracy can not be guaranteed.
In this paper, an easy-to-use generic positioning framework is proposed, which allows users to plug in a single or multiple positioning algorithms. We illustrate the usability of the framework by discussing a new hybrid positioning solution. The combination of a weighted (range-based) and proximity (range-free) algorithm is made. Both solutions separately have an average error distance of 13.5m and 2.5m respectively. The latter result is quite accurate due to the fact that our testbeds are not sparse distributed. Our hybrid algorithm has an average error distance of 2.66m only using a selected set of nodes, simulating a sparse distributed sensor network. All our experiments have been executed in the iMinds testbed: namely at “de Zuiderpoort”. These algorithms are also deployed in two real-life environments: “De Vooruit” and “De Vijvers”.

A.1 Introduction

Combining wireless sensor network nodes with the upcoming trend of smartphones creates a totally new range of possibilities. Normally, wireless sensor networks are used to monitor a certain environment and measure e.g. the temperature and humidity. But also tracking of persons and equipment can be done by sensor networks. GPS [1] is the traditional way of tracking people or vehicles outdoor, however this does not work properly indoor because line of sight (LOS) is required to receive the GPS signals.

Sensor nodes inside buildings can fix this issue, however other factors have to be taken into account: interference, infrastructure, the amount of sensor nodes that is required, energy consumption,... It will always be a trade-off between cost and accuracy. Further, a myriad of positioning algorithms have been developed in the last few years. A standalone solution generally does not offer sufficient accuracy in different environments (indoor/outdoor, different type of buildings,...). In this paper however, we will try to find a solution with an acceptable accuracy when only a sparse distributed sensor network is available. Our algorithm described in this work is a combination of two already existing algorithms. Each belonging to a different subdivision, namely range-free and range-based. Both solutions show too many defects in thinner environments. Combining them results in a noticeable improvement. In this way, room accuracy can be guaranteed.

The rest of the paper is organized as follows. In Section A.2, the already existing algorithms are clarified. Section A.3 describes our generic architecture framework. The hybrid algorithm build in this framework whereby the two previous are combined is discussed in Section A.4. Section A.5 handles about our different testbeds. The experiments with their results are summarized in Section A.6. Finally, some conclusions can be made. These are, together with the future work, clarified in Section A.7.

A.2 Related work

In this section, we conclude the work that is essential to comprehend our hybrid solution. Localization algorithms can be subdivided in two different categories. The first category is called the "range-based"-algorithms. In order to calculate a position pertaining to multiple fixed nodes, a distance measurement is essential. Then, on the base of this distance, a position can be determined by means of trilateration. Trilateration is a method to find the intersection of three circles whose center and radius are known. There are many different ways to measure the distance. The most familiar techniques are Received Signal Strength Indication (RSSI), Time of Arrival (ToA), Time Difference of Arrival (TDoA) and Angle of Arrival (AoA).

The second category, "range-free"-algorithms, does not require a distance measurement to calculate the position of a sensor node. They are based on the information of the connection. If two sensor nodes can connect to each other, than the maximum distance between them is the maximum transmission range. Thus the position of the mobile node can be estimated with this information. This is a very simple and cheap technique. Moreover the accuracy will depend on the density of the wireless sensor network. Centroid, triangle elimination and proximity are common range-free algorithms.

The hybrid solution uses both techniques. A combination of a range-based and a range-free algorithm is made. In the following two sections, both algorithms will be explained more in detail.

A.2.1 Range-based: weighted

The first one is a range-based solution described by Tareq Ali Alhmiedat et al. in [2]. The proposed algorithm is based on weighted RSSI values. The main idea of RSSI is that the transmission power P_T directly affect the received power P_R of a signal. Via the Friis transmission equation, also defined in [2], the linear relationship becomes clear:

$$P_R = P_T * G_T * G_R \left(\frac{\lambda}{4\pi d} \right)^2 \tag{A.1}$$

where G_T, G_R are gain of transmitter and receiver respectively. λ is the wavelength of the signal and d is the distance between sender and receiver. The received signal strength indicator (RSSI) can be defined as the ratio of the received power to the reference power P_{Ref}.

$$RSSI = 10 * \log \frac{P_R}{P_{Ref}} \quad \text{(A.2)}$$

Each RSSI value can be matched with a certain distance. The proposed algorithm in [2] not only uses the RSSI-values to measure the distance between a fixed and mobile node, but also the distance between the fixed nodes mutually is measured. These values function as weight factors for the distance calculation between the fixed and mobile node. These weight factors are shown in Figure A.1 as $w[htbp]12$, w_{13} and w_{23}. The distance from M to, for example, B_1 can be calculated as follows:

$$Distance(M, B_1) = \frac{RSSI\, B_1 * w_{12} + RSSI\, B_1 * w_{13}}{2} \quad \text{(A.3)}$$

Their results prove that these weight factors add value to the accuracy. A drawback of the RSSI technique is that these measurements are very sensitive to the environment and potential changes in it. The relationship between the distance and RSSI depends on the room. For example, in a long corridor, the fixed nodes their signals will have a greater range because they reverberate through the long walls. In this way, completely different results can be obtained.

A.2.2 Range-free: proximity

In contrast to the previous category, range-free algorithms do not take RSSI-values into account. If a mobile sensor node has a range of 10 meters, than a fixed node can only receive his messages if the mobile node is maximum 10 meters away. This is the only information that is used to calculate the position of a mobile node.

This technique is used by J. Wyffels et al. in [3]. A proximity-based algorithm is used to localize the patients and the nurses in a healthcare environment. Important here is that the transmission power is well configured. If the power is too low, the mobile node could be out of range between two fixed nodes. And also vice versa if the power is too high, too many fixed nodes will receive the beacon and a wrong estimation could be made.

The latter problem can be solved by using a centroid algorithm [4]. This is only useful if there is a set of fixed nodes with an overlapping coverage area. The beacon of the mobile node is received by multiple fixed nodes. In order to determine the position, the centroid of all the receiving fixed nodes is calculated (Eq. A.4).

$$[x_M, y_M] = \left[\frac{\sum_{n=0}^{k-1} x_n}{k}, \frac{\sum_{n=0}^{k-1} y_n}{k} \right] \quad \text{(A.4)}$$

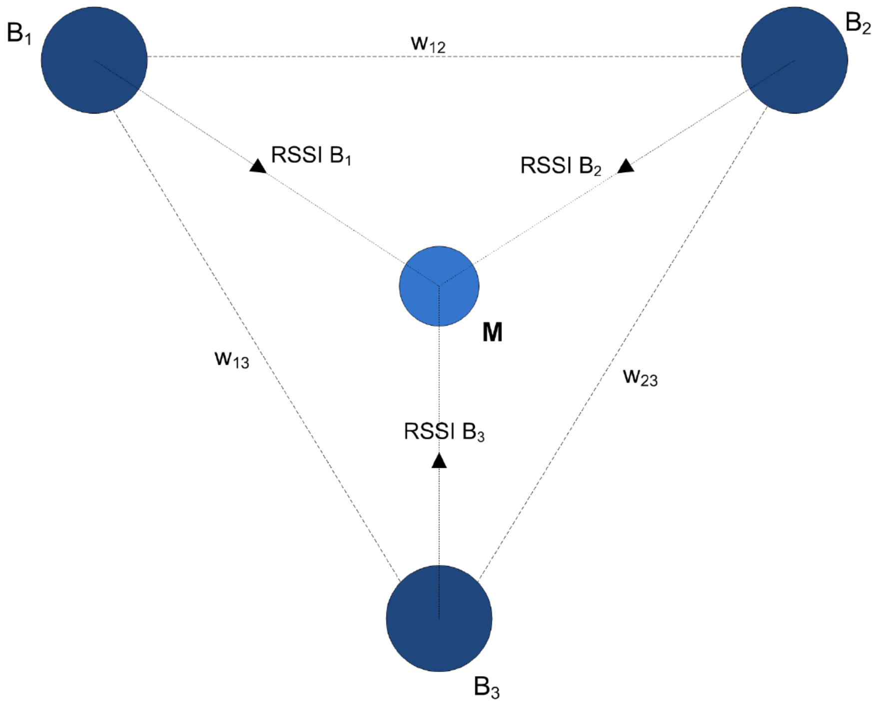


Figure A.1: The graphical scheme of the Weighted algorithm. Circles B_1, B_2 and B_3 represent the anchors whilst M acts as the mobile node. $RSSI\ B_1$, $RSSI\ B_2$ and $RSSI\ B_3$ are the RSSI values between the mobile and the anchor nodes whilst w_{13}, w_{12} and w_{23} are the RSSI values between the anchor nodes mutual.

Normally would this algorithm give a 100% guarantee that room-accuracy is ensured. However, experiments have shown that this is not always the case. If the walls are small enough and/or not made of concrete, signals can go trough and a fixed node in a different room can catch up the beacon. In order to prevent a wrong location estimation, some extra logic can be implemented in the algorithm.

To implement the extra logic, some extra information is necessary as well. Suppose we have the exact coordinates of all the walls, doors and nodes inside a building. Knowing that every beacon has an index number, the direct path could be checked between the two fixed nodes who received the consecutive beacons. If the mobile node goes from one room to another, without using a door, then the last beacon can be dismissed. For example (Fig. A.2) when node A_2 receives a beacon and the next beacon is received by node B_2. It is impossible to move directly from A_2 to B_2 without passing nodes A_1 and B_1. So the message that was received by beacon B_2 will be rejected.

With this optimization room-accuracy can be guaranteed. Still, this solution has the drawback that a lot of fixed sensor nodes are necessary to retrieve good results. If the network is sparse distributed, then the algorithm would not work properly.

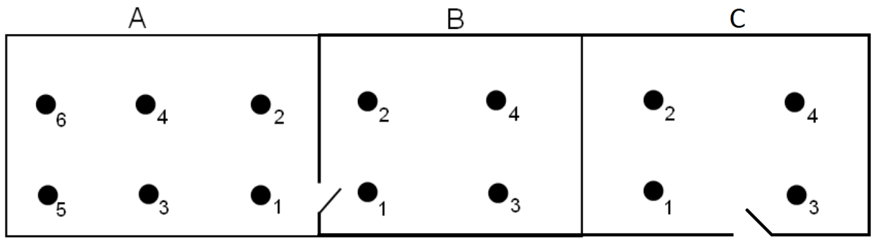


Figure A.2: An example of three neighbouring offices. Walls and doors help the solution to improve the accuracy based on normal logic: people cannot walk through walls.

A.3 Positioning framework

The framework is developed in Java and consists of three parts: the positioning server, the web server and the client application (Fig. A.3).

The positioning server has two functional blocks. The interconnection gateway is responsible for the retrieval of positioning information gathered by the network infrastructure or mobile unit that is being located. The interconnection gateway further incorporates an abstraction layer which hides the underlying technology (ZigBee, Wi-Fi, Bluetooth, ...) from the positioning server. In Figure A.3, two different approaches for positioning in wireless sensor networks are shown. On the left side, a mobile device

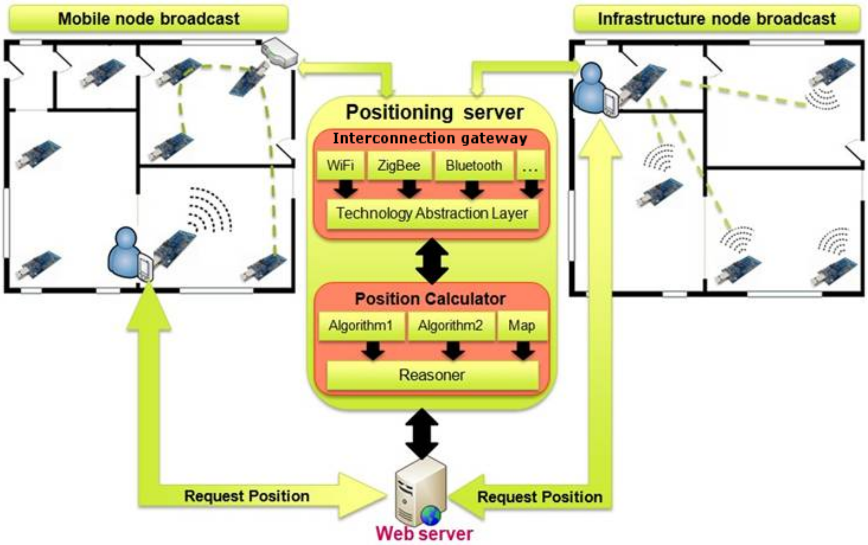


Figure A.3: Schematic overview of the framework architecture: two types can be differentiated, mobile or infrastructure node broadcast. Both can be connect with the positioning server which consists of an interconnection gateway and a position calculator.

broadcasts positioning beacons and the sink node of the WSN forwards the beacons to the interconnection gateway. On the right side, the infrastructure nodes broadcast beacons and the mobile unit collects and forwards the beacons to the interconnection gateway. The interconnection gateway further passes the positioning information to the position calculator, which consists of pluggable positioning algorithms. Multiple positioning algorithms can be active at the same time. A reasoner is used to select the algorithm giving the most accurate position or to intelligently combine the results of multiple algorithms into a more accurate (hybrid) position. Map info can also be taken into account when calculating the position.

The web server can poll the positioning server for the user's position. And the client application can either run on a smartphone or a central monitoring station. The client communicates with the web server through e.g. Wi-Fi or Ethernet.

Some advantages of the framework:

- Existing smartphone applications can use position information by implementing a simple interface allowing the application to request a user's position from the web server.

- Conversion of relative coordinates to GPS notation is possible. This implies that client applications developed to work outdoor (GPS), can easily use this framework.

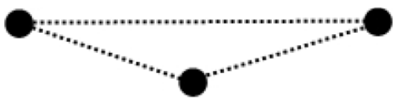
- The user of the client application can pinpoint his correct location on the floor plan (for testing purposes). The application then calculates the difference between the estimated and the real position, thus allowing the user to evaluate the algorithm.

A.4 Hybrid algorithm

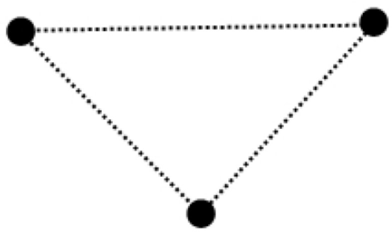
Having this framework described above, designing a hybrid solution is very efficient. The reasoner allows the position calculator to combine the results of different algorithms and other available information. In the hybrid solution the reasoner has two choices: if the mobile node is in range of a fixed node we use the result of the proximity algorithm. If no fixed node can hear the proximity node, the reasoner will decide to use the weighted RSSI algorithm, where the mobile unit has a wider range.

The biggest problem of the stand-alone weighted algorithm, is the selection of the nodes. An ideal situation would be that the proximity node would be surrounded by the closest fixed nodes. As discussed in subsection A.2.1, selecting the closest nodes is not possible if only the RSSI values are available. The proximity algorithm can give extra information whereby finding the closest nodes can be realized.

Hence, the node selection can be optimized using the latest information of the proximity algorithm. In that way, the first node of the triangle is determined. In order to have a good coverage of the area, the two other nodes must be well selected. If the angles of the triangle are too sharp (Fig. A.4a) than the weighted algorithm will not function properly. In certain situations, the two last nodes will have to be reselected until a good triangle (Fig. A.4b) is founded.



(a) Example of a bad triangle



(b) Example of a good triangle

Figure A.4: Two types of triangles, a good and bad example. The goals is to select nodes which resemble the good triangle.

Once the three fixed nodes are selected, a distance measurement is the next step in the procedure. This is done the same way as the stand-alone weighted algorithm (Eq. A.3) except for one thing. The RSSI values are

Data: Three circles of each fixed node
Result: Position of the mobile node
if *three circles do not intersect* **then**
  while *smallest circle does not intersect with the second smallest circle* **do**
    increase the smallest circle
  end
end
`// Now at least two circles intersect`
Calculate the intersection of the two smallest circles
position mobile node = intersection of the two smallest circles closest with the biggest circle

Algorithm 1: Adapted trilateration

slightly adapted because results of previous experiments have shown that the calculated distance was almost always too big. This adjustment is estimated experimentally. After the distance calculation, the three circles can be created and trilateration can be applied. In perfect circumstances, the three circles will intersect in exactly one point. However, in practice this is never the case. Due to the environment and interference, the three circles will never intersect in one single point. Therefore, an adapted trilateration technique is shown in Algorithm 1.

Finally, if the reasoner has access to other input, such as information about walls, rooms, doors, we can use this to influence our position estimate.

A.5 Environment descriptions

This positioning framework including the hybrid solutions has been tested in two wireless testbeds and also in two different real-life situations. Each environment will be explained further in detail in the next sections.

A.5.1 w.iLab-t at the "Zuiderpoort"

The w.iLab-t is an extensive facility that is introduced in detail by S. Bouckaert et al. in [5]. The infrastructure is distributed on three floors of the iMinds office in Ghent, Belgium (Fig. A.5). The network consists of 200 nodes. Each node has (i) a Tmote Sky IEEE 802.15.4 mote, (ii) two Compex WLM54SAG 200mW AR5006XS 802.11a/b/g 54/108 Mbps miniPCI wireless cards and (iii) an environment emulator. The latter one is self-made and used for simulations: environment (e.g. temperature change),

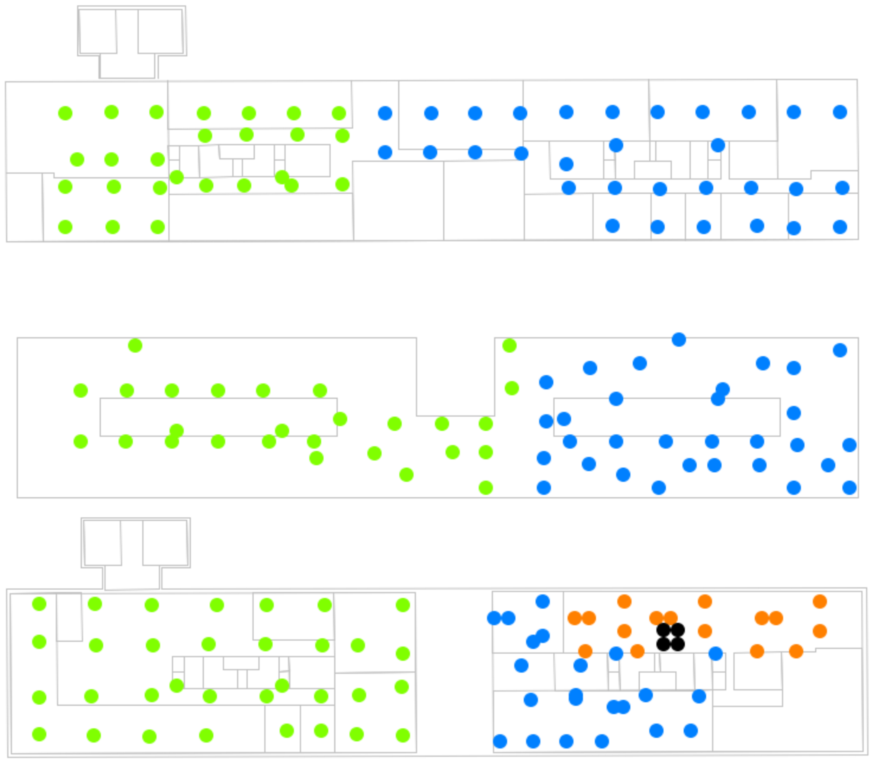


Figure A.5: w-iLab.t at the "Zuiderpoort". 200 nodes on three floors (18m x 90m). The different colors indicate the nodes can be divided in groups for executing tests.

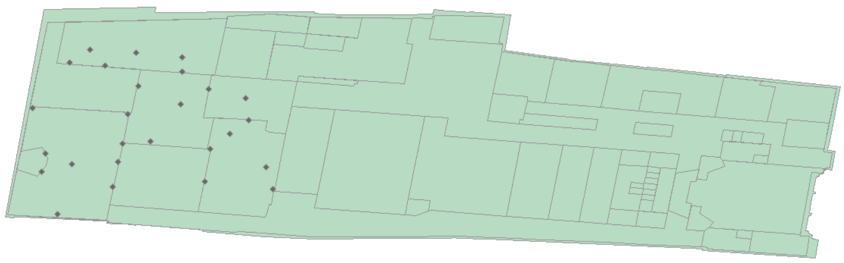
battery drop, user input, etc.

These nodes are centrally managed for control and monitoring purposes and remote access by using an Intel x86 architecture (PC Engines Alix Boards).

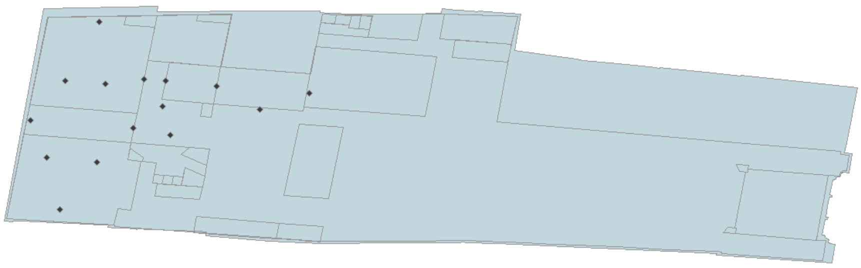
A.5.2 "De Vooruit"

De Vooruit is an ancient building close to the historical centre of Ghent. In the past, this building was a place for the working class where they could eat, drink and enjoy culture at democratic prices. Since 1982 De Vooruit is recognized as a monument and nowadays it is still used to organize lectures, debates, concerts, parties, ...

This location was a perfect use case to test the indoor localization solutions. Due to the fact that the building was recognized as a monument, it was not allowed to use a cabled network. In this situation, wireless sensor networks were the only solution to handle this problem. 50 nodes, distributed over four different floors (Fig. A.6), were used to locate the mobile nodes worn by the visitors. In this use case, Sentilla JCreate nodes in combination with battery packs were used.



(a) Fourth floor of De Vooruit



(b) Sixth floor of De Vooruit

Figure A.6: Floorplan of the fourth and sixth floor of De Vooruit, the diamonds represent the fixed nodes installed in the building.

A.5.3 "De Vijvers"

As a second use case, the positioning was tested in a home for the elderly. The goal here was to track people with dementia that are not allowed to leave the home. When a person goes in a restricted zone, an alarm was sounded. The position of the person could then be seen on a map in the reception. In this building (Fig. A.7), 25 Sentilla JCreate nodes were attached to obtain the required accuracy for this application.

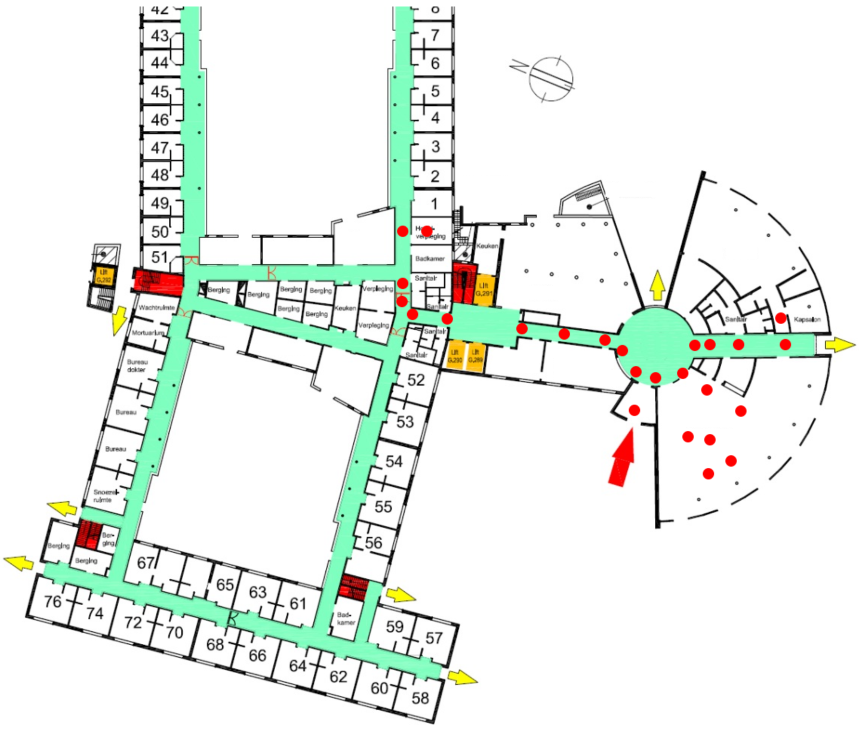


Figure A.7: The southern part of "De Vijvers" where 25 nodes were attached in the central and eastern part of the building. Their positions are marked with red dots.

A.6 Results

In this section, we present the results of all the interesting measurements. First, the two algorithms are tested separately, followed by the results of the hybrid solution. All these measurements are done at De Zuiderpoort (Section A.5.1) on the third floor.

A.6.1 Range-based: weighted

The results from [2] showed that the weighted RSSI-values give a more accurate position than the normal RSSI-values. For those reasons, only the results of the weighted algorithm will be shown.

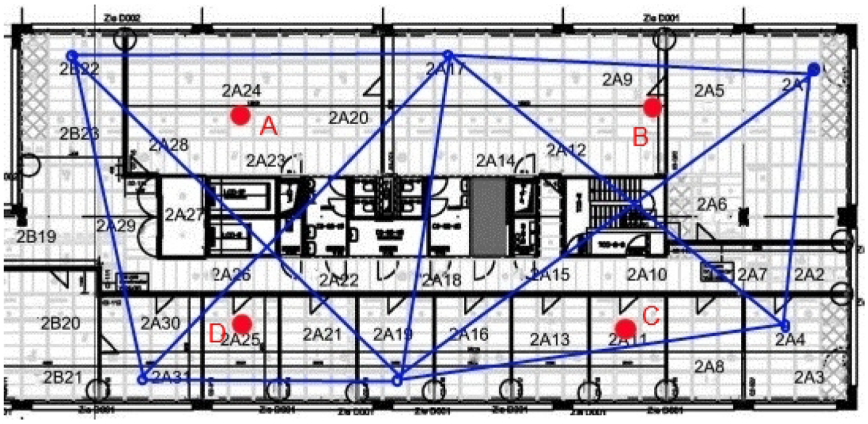


Figure A.8: Dividing the third floor in big triangles for the second test of the weighted algorithm.

Several tests have been executed. First, all available nodes on the third floor were used. This, however, gave very poor results so they are not published in this work. Some measurements had an error distance of more than 20 meters. An explanation for these large error distances is multipath fading of the nodes in the corridor. The setup of the second test is shown in Figure A.8. The third floor is divided in big triangles (marked with blue lines) to calculate the position of A, B, C and D (marked with a red dot). The results of these measurements can be found in Figure A.9. For each location, ten measurements are executed. The smallest error distance is $6.3m$, the biggest is $21.2m$ with an average error distance of $13.8m$. These results are not acceptable because room accuracy cannot be guaranteed. The large error distance is due to the fact that a high transmission power was necessary to communicate trough the concrete walls in the center of the building. The concrete walls has a strong influence on the RSSI measurements. For those reasons, a third test was implemented that avoids the concrete walls.

The triangles of the third test can be found in Figure A.10. In this way, the signals do not need to go through the concrete walls so a lower tx-power can be used. The results of this setup are represented in Figure A.11. Again, in this test, ten measurements at each location are recorded. The smallest error distance here is only $0.5m$, the largest one $8.2m$. The average error distance is $4.8m$. These results are much better than with the large triangles, but still, an error distance of $8m$ is unacceptable.

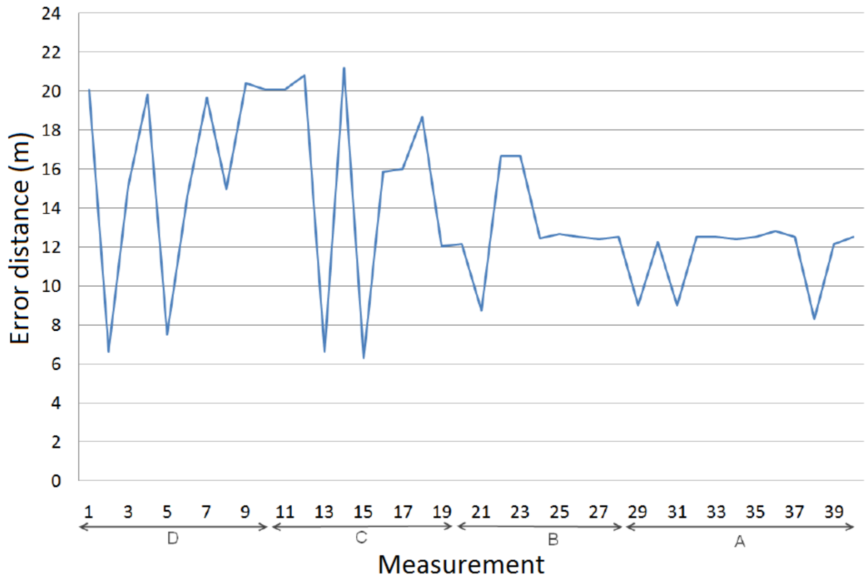


Figure A.9: Graphical overview of the results of the weighted algorithm using a sparse distribution of the fixed nodes

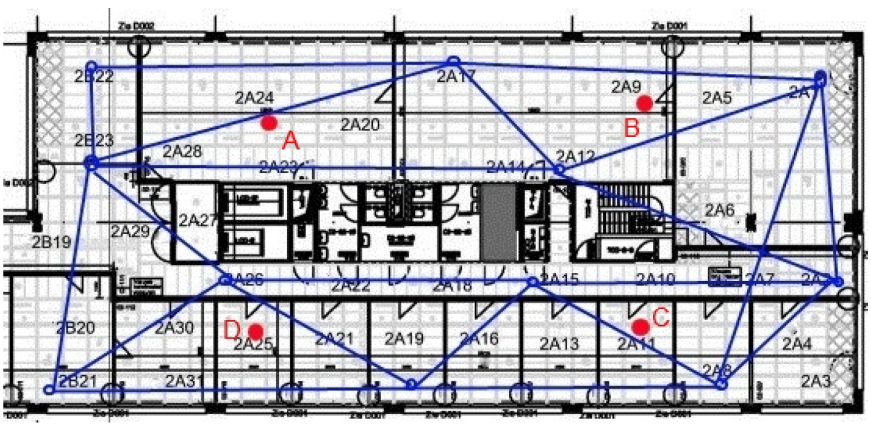


Figure A.10: Dividing the third floor in small triangles for the third test of the weighted algorithm.

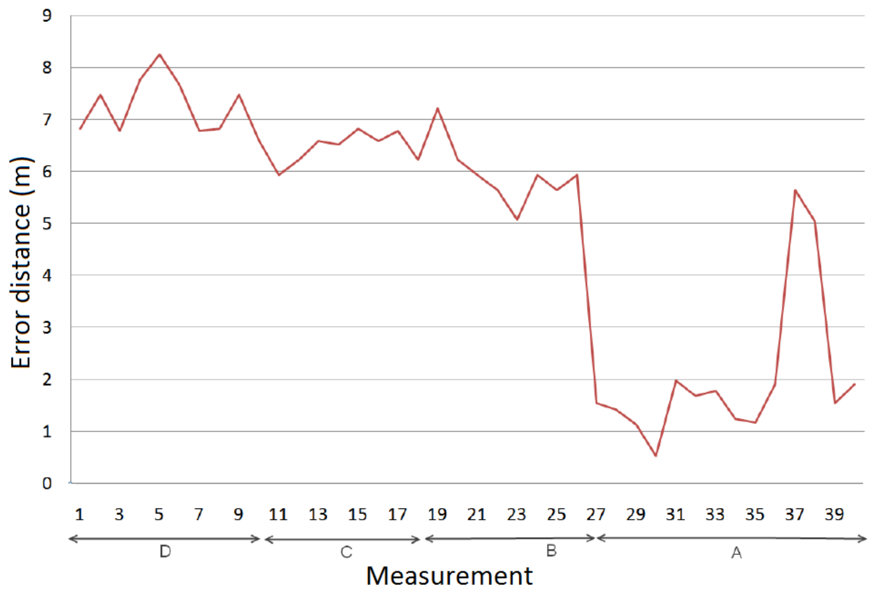


Figure A.11: Graphical overview of the results of the weighted algorithm using a dens distribution of the fixed nodes

In these results, it became clear that this single algorithm was not capable to deliver the room accuracy.

A.6.2 Range-free: proximity

The results of the proximity based algorithm are completely dependent of the used infrastructure. The density of the fixed nodes determines the accuracy of the localization. Our algorithm is tested in the w-iLab.t at De Zuiderpoort (Section A.5.1) where the fixed nodes have an intermediate distance of 5 meters. This means that the maximum error distance is about $2.5m$. In the best case, the mobile node is located right under the fixed node, meaning that the error distance is $0m$.

A.6.3 Hybrid algorithm

The hybrid algorithm is designed to work properly in sparse distributed sensor network. However, the w-iLab.t is not sparse distributed. For this reason, it was hard to retrieve results using the whole testbed, the proximity beacons were always reachable whereby the hybrid algorithm almost never switched from the proximity to the weighted algorithm. This produced the same results as in Subsection A.6.2. In order to test this algorithm its full functionality, some artificial tests are done. From the moment a fixed node received a proximity beacon, the transmitting of proximity beacons by the mobile node will be stopped. Hereby, the weighted algorithm has to come active to calculate the final location of the node.

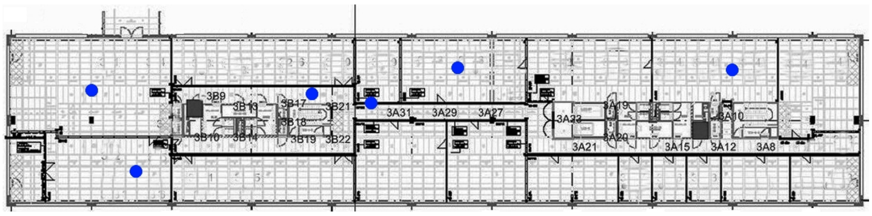


Figure A.12: Positions of the mobile node for testing the hybrid localization solution

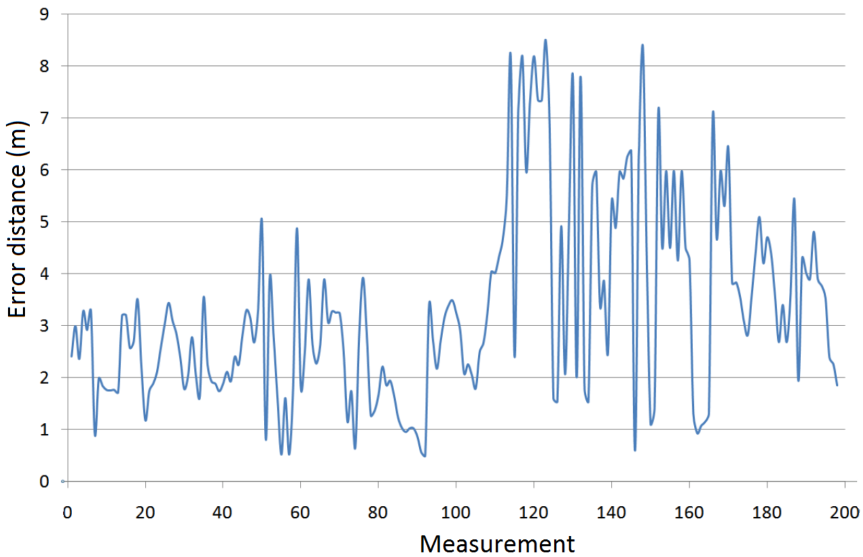


Figure A.13: Graphical overview of the results of the hybrid algorithm using a sparse distribution of the fixed nodes

The mobile node is placed at different locations on the third floor in De Zuiderpoort building, these are marked in Figure A.12 with the blue spots. The results of these measurements can be found in Figure A.13, these are the worst possible results because the proximity is often disabled in order to activate the weighted part of the hybrid algorithm. 200 measurements were made across the different locations. The minimum error distance was $0.49m$ and maximum $8.5m$. The average of all the measurements together was $3.28m$. The worst results are due to the fact that some fixed nodes are placed in ventilation ducts. These are hard to reach for the signals of

Table A.1: Summary of the experimental results

LOCALIZATION SOLUTION	ERROR DISTANCE (IN METER) MINIMUM	MAXIMUM	AVERAGE
Weighted algorithm *(big triangles)*	6.3	21.2	13.8
Weighted algorithm *(small triangles)*	0.5	8.2	4.8
Proximity algorithm	0	2.5	-
Hybrid algorithm *(all nodes)*	0.49	8.5	3.28
Hybrid algorithm *(filtered nodes)*	0.49	8.5	2.66

the mobile node. The RSSI-values of these messages are extremely low causing a greater error on the distance calculations from the mobile node to the fixed node in the ventilation duct. This affected the results significantly, when we drop all the results of the fixed nodes in the ventilation ducts, the new average error distance is $2.66m$.

Hence, this algorithm has also some drawbacks. Each algorithm uses a different transmission power. It is very important that the proximity algorithm his transmission range can be limited to the half of the distance between the fixed nodes. The idea is that only one fixed node can receive the beacons at a time. But with the weighted algorithm, enough nodes need to receive the beacons from the mobile node in order to make triangulation work properly. The tx power of a Tmote Sky can be programmed dynamically, but in our case, extra attenuators were necessary to reduce the transmission range. To fix this issue in our situation, two mobile nodes were used.

A.6.4 Summary

A summary of all the experimental results can be found in Table A.1, all the minimum, maximum and average error distances are collected in one organized table. It becomes clear that the hybrid solution has an improvement (if you compare the average error distances).

A.7 Conclusion and future work

This paper presents a hybrid indoor localization solution designed to achieve room accuracy using sparse distributed sensor networks. Hereby a positioning framework is developed to accomplish a hybrid solution, based on two existing solutions.

The positioning framework consists of two functional blocks: the interconnection gateway and the position calculator. The interconnection gateway gathers all the necessary data from the fixed and mobile node in order to calculate the position. This data can come from any kind of hardware device/technology. The position calculator contains all the different localization algorithms that calculate the position of the mobile node using the data from the interconnection gateway. This calculator includes also a reasoner that decides which algorithm calculates the most accurate position at a certain moment, it can also combine multiple algorithms to improve the accuracy even more.

The hybrid solution is based on a range-based and range-free algorithm. The former is a category of techniques that requires distances measurements in order to calculate the position of a mobile node. These distance measurements can be done in different ways. In this paper, the range-based "weighted" algorithm is proposed. It uses RSSI measurements to calculate the distance between the nodes. The higher the value, the shorter is the distance between the nodes. Innovative here is that RSSI measurements are also used to calculate the distance between the fixed nodes mutually. Using this extra information, a *weighted* distance calculation can be done using triangulation.

The range-free solution "proximity" does not require these distance calculations, the localization is only based on the information of the connection. This means that the reception range of a fixed node is as well as the maximum error distance. However, an extra optimization is possible, if multiple fixed nodes receive a beacon, then the centroid of all the fixed nodes can be calculated and be assumed as the point closest to the mobile device.

Both algorithms show issues in sparse distributed sensor networks. The accuracy of the weighted algorithm is far from acceptable because it is not easy to determine the correct triangle for the calculation and the proximity solution is completely depended on the density of the fixed nodes. Therefore, combining both algorithms can resolve the biggest issues of both solutions. First a proximity beacon is received by a fixed node, this is the first corner of the triangle. Then the other two corners are determined in order to get a good triangle.

In the results, it became clear that the improvement of the hybrid solution is significantly. The average error distance dropped from $13.8m/4.8m$ to $3.28m/2.66m$. Still, some future work can be done. First, the issue with the transmission power must be tackled. Further, comparative tests using WiFi or other technologies are in progress.

References

[1] N. Bulusu, J. Heidemann, and D. Estrin. *GPS-less low-cost outdoor localization for very small devices*. Personal Communications, IEEE, 7(5):28–34, 2000.

[2] T. A. Alhmiedat and S.-H. Yang. *A ZigBee-based mobile tracking system through wireless sensor networks*. International Journal of Advanced Mechatronic Systems, 1(1):63–70, 2008.

[3] J. Wyffels, J.-P. Goemaere, P. Verhoeve, P. Crombez, B. Nauwelaers, and L. De Strycker. *A novel indoor localization system for healthcare environments*. In Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on, pages 1–6. IEEE, 2012.

[4] T. Qiu, Y. Zhou, F. Xia, N. Jin, and L. Feng. *A localization strategy based on n-times trilateral centroid with weight*. International Journal of Communication Systems, 25(9):1160–1177, 2012.

[5] S. Bouckaert, W. Vandenberghe, B. Jooris, I. Moerman, and P. Demeester. *The w-iLab. t testbed*. In Testbeds and Research Infrastructures. Development of Networks and Communities, pages 145–154. Springer, 2010.